Symmetry and Reformulation:
On Intellectual Progress in Science and Mathematics

by

Joshua Robert Hunt

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
in the University of Michigan
2022

Doctoral Committee:

Professor Gordon Belot, Co-Chair
Professor Laura Ruetsche, Co-Chair
Associate Professor David John Baker
Professor Henriette Elvang

Gustave Doré, “Deer in a Pine Forest” (ca. 1865)

Joshua Robert Hunt

joshhunt@umich.edu

ORCID iD: 0000-0001-5150-0388

Dedicated to Ken Manders, for opening the door,
and to Jeremy Butterfield, for keeping it wide open!
Thank you for wearing your passion *on your sleeves*!

Acknowledgments

In both the thrill of the chase and the tortuous assessment of how it all hangs together, I have been helped by many—many more than I can thank here! Undoubtedly, this dissertation has been the most intellectually gratifying experience of my life. Yet, I am surely in good company to admit that it has at times felt somewhat alienating. Over the past five years, it has often seemed unclear how I ought to be spending my time, or what my ultimate goals should be. Hopefully, the more uncomfortable parts of this experience have taught me a number of lessons and skills. I would like to think that I have learned how to sustain myself intellectually on the basis of questions and dialectics that I find intrinsically interesting. I have acquired a sense of pure self-motivation and a dogged determination to keep going. If I were left alone in Siberia with nothing but my laptop and sufficient provisions, I would keep writing!

I thank my committee for affording me supreme and unfettered freedom in pursuing my intellectual dreams. Only time will tell whether or not this freedom was for the best, although it is most definitely what I wanted. I especially thank Professor Belot for his sustained and detailed comments on various chapters. The quality of these comments gleamed with a sense of the supererogatory, and for that I am very grateful. I thank Professor Elvang for inspiring me to look into examples from quantum field theory. These examples dramatically increased my confidence that compatible reformulations are a central aspect of contemporary physics. I thank Professors Ruetsche and Baker for some highly orienting conversations, comments, and reading suggestions for various chapters. My committee has also generously provided a great deal of help toward projects not represented in this dissertation. I thank them in advance for what I imagine will be extremely helpful—and hopefully not too devastating—comments!

More than anyone, I thank Elise Woodard for countless conversations and constructive suggestions regarding most of the issues and arguments in this thesis. She has shown

me out of many a fly bottle, along with reassuring me that I've escaped a few on my own accord. I know that she would rather me be not so unsettled! Elise has read and commented on almost everything here, and has had an ample hand in gently shaping these ideas over the years. Her own intellectual grit, scholarly style, professional exuberance, and organizational acumen has been my single greatest source of inspiration for the past six years. If not for her love and support, I might not have finished this dissertation.

For validating many aspects of my meta-philosophical temperament, I thank Cameron McCulloch. I have met very few people who viscerally feel what I feel about philosophy (both in theory and in practice). I also heartily thank him for becoming a minimalist (of the material variety!) and reselling a phenomenal collection of OUP books at barn-buster prices. There are many sources I might not have engaged with if not for his curiosity and generosity. Both Cameron and Elise deserve credit for inspiring me to write on a regular basis. Much of this thesis was written as part of a weekly writing competition with Elise, using techniques advocated by Cameron. Hopefully it's now my turn to inspire them!

I thank Arvind Prasadan for early encouragement that the topic of this thesis has wide applicability, along with relaying examples in electrical engineering, statistics, and applied mathematics. If he ever reads this, I hope he'll see that I really have made progress on these questions since our undergraduate days! I thank Joe Andros for relaying examples in financial mathematics, Carsen Stringer for examples dealing with neural nets and machine learning, and Alex Kavner for some examples in physics. I would like to have incorporated these multifarious examples, but alas! It is work for another day.

Jonathan Sarnoff is responsible for motivating me in the fall of 2020 to look into metaethics, which I found deeply illuminating. I owe him immensely for this, since otherwise I would have kept struggling to express key portions of my positive account. Both Abdul Ansari and Adam Waggoner provided encouragement on this front as well. I thank Professor Gibbard for some orienting exchanges, along with his body of work on expressivism (which was life-changing). Better late than never, as they say! Dovetailing with my new-found passion, Professor Maegan Fairchild ran an *incredible* Winter 2021 reading group on Amie Thomasson's truly delightful *Norms and Necessity* (2020). This group and book gave me the courage to apply similar ideas to some problems in philosophy of science that had tormented me for years.

I thank my family for loving me near-unconditionally. I know that it is not the path they would have chosen for me. Their support throughout graduate school has been invaluable, through the good times and the tough. To my parents especially: thank you for helping me live the life I imagined. I have been greatly humbled by the support and complete acceptance my family has shown me through periods of personal trauma. They have moved me to tears. Thank you, Emily. Thank you, Sean.[1]

For additional pastoral care, I thank Dr. Howard Schubiner, whose good cheer and advice did more to alleviate my chronic hand-pain than anything else in the past 14 years. To any contemporaries reading this suffering from chronic, computer-related pain, please feel free to get in touch with me.

I would like to especially thank Dr. Jeremy Butterfield and Prof. Bryan Roberts for graciously organizing their "Philosophy of Physics Boot Camp" on Zoom throughout the pandemic. Before joining what we might call "The Cambridge Circle," I had no idea how invigorating and inspiring a philosophy of physics reading group could be. Much of my reason for continuing to work within philosophy of physics stems from the inspiration derived from this reading group. It has truly soothed my intellectual soul. Jeremy in particular played a key role in keeping my spirits up throughout some trying stints on the academic job market. Along with Prof. Ken Manders, he is a paradigm of the kind of mentor I hope to be someday.

Much of this dissertation was written while imagining what Manders might say, based on extensive and extremely formative conversations at the University of Pittsburgh from 2012 to 2014. I thank Ken and Weia for hosting me in 2017 and 2019 for some very fruitful discussions on the early stage directions of this project. I didn't work up the courage to send him enough of it for comments, and I'm not sure he would endorse the direction that it's taken. He probably would have tried to talk me out of a lot of it. But I hope he won't think I've strayed *too far* from the right path. Even all these years later, it still feels that almost everything I learned about doing philosophy, I learned from Manders.

Staying on the seam of unbridled intellectual passion, I would like to thank Gabriele Carcassi, who in the Fall of 2020 welcomed me into his research program just as my spirits were drooping. His generosity with his time is a model of interdisciplinary exchange.

[1]Truth be told, it is really Sean's fault for getting me into this whole mess. He gave me in 2007 a copy of Will Durant's *The Story of Philosophy*, which gave voice to my philosophical temperament.

Likewise, I thank Prof. Christine Aidala for both research support and her contributions to their ongoing Assumptions of Physics project. Although that research program is not reflected here, it should be clear that this project manifests a kindred spirit.

Over the years, the philosophy graduate students of the University of Michigan have provided sage wisdom on many parts of this project. This includes two talks on material in Chapters 1 and Chapter 4 (in January and October 2017, respectively) and a March 2021 talk on Chapter 5. They have seen two versions of Chapter 6 (in February 2020 and October 2021). Both times, I viscerally felt that I had figured out what it means to make a property manifest, only to discover strong grounds for dissatisfaction. In particular (with a real risk of embarrassing omission!), I thank Abdul Ansari, Kevin Blackwell, Jason Byas, Francisco Calderón, Kevin Craven, Guus Duindam, (the late) Ian Fishback, Paul de Font-Reaulx, Gillian Gray, Emma Hardy, Rebecca Harrison, Malte Hendrickx, Jesse Holloway, Gabrielle Kerbel, Jaewon Kim, AJ Kuhr, Brendan Mooney, Chris Nicholson, Sumeet Patwardhan (for consistently incisive comments!), Caroline Perry, Ariana Peruzzi, Josh Petersen, Laura Soter, Alvaro Sottil de Aguinaga, Angela Sun, Adam Waggoner, Sophia Wushanley, and Glenn Zhou. For advice intellectual and practical, I thank Sara Aronowitz, Mercy Corredor, Daniel Drucker, Anna Edmonds, Zoë Johnson King, Sydney Keough, Eli Lichtenstein, Eduardo Martinez, Filipa Melo Lopes, Cat Saint-Croix, Jonathan Sarnoff, Chip Sebens, Umer Shaikh, Rohan Sud, and Nina Windgätter. I couldn't imagine better near-peer role models. In their capacity as visitors, Maria Altepeter, Claire Field, Nikhil Venkatesh, Reza Hadisi, and Alex Meehan all provided excellent comments on some of these talks. For getting Elise and me through the pandemic, I thank Calum McNamara and Maria Rosala. For crucial life-support this year, I thank Sean Costello. For embodying the fact that there's more to life than philosophy, I thank Johann Hariman.

Many University of Michigan faculty provided constructive comments on versions of Chapters 2 and 4, presented as 'Brown Bags' in the Fall of 2019 and 2020. I especially thank Professors Sarah Buss, Maegan Fairchild, Jim Joyce, Sarah Moss, Tad Schmaltz, Eric Swanson, Jamie Tappenden, and Brian Weatherson. I thank Woodard's Epistemology Work in Progress group (E-WiP) for sponsoring a March 2020 talk on epistemic vs. practical normativity and especially Alex Meehan for extended conversations (right before COVID lockdowns began). This helped me clarify the general framework in Chapter 2. I thank Professor Stephan Hartmann for catching a mistake in one of my examples in Chapter 1.

A version of Chapter 3 was presented at the 2021 Philosophy of Science Association conference in Baltimore. Thanks especially to Prof. Kareem Khalifa and Prof. Ruetsche for detailed comments. I thank Prof. Hasok Chang for advising my earliest work on scientific understanding at the University of Cambridge, and for comments and encouragement on the dissertation manuscript. I hope to put Hasok's advice *into action* in future work!

I presented versions of Chapter 4 at the 2017 Midwest Philosophy of Mathematics workshop (Notre Dame), the 2018 MuST Models of Explanation conference (Turin), the 2019 LMP graduate conference (Western University), and the 2019 Cushing Workshop (Notre Dame). The initial case study goes back to 2014, and was presented at the Philosophy of Chemistry symposium at the PSA in Chicago. I thank all of the audience members who asked questions, and especially Professors Ruetsche and Belot for extended feedback during a 2017 directed reading course. Prof. Jamie Tappenden provided feedback in his 2016 seminar on causal and non-causal explanation. Patrick Manzanares and David Schroeren provided extensive and constructive comments. I thank Prof. Michael Strevens for a series of helpful conversations in 2019 while visiting NYU.

More recently, I presented parts of Chapter 6 at the 2022 Philosophy of High Energy Physics workshop at the University of Pittsburgh. More than anything, this confirmed my suspicions that philosophers of science remain wary of normativity (at their peril!). I thank Prof. Porter Williams for advice on philosophical aspects of quantum field theory while visiting USC in the late fall of 2019. Prof. James Wells gave some orienting advice and helpful reading recommendations on the topic of gauge choices (I thank Prof. Wells as well for being a stalwart contributor to our philosophy of physics reading group!). I thank Angela Sun and Carolina Flores for pointing me toward some extremely relevant sources for Chapter 5. Prof. Baker suggested some literature that would have dramatically improved this chapter, but these will have to await a future project!

The expressivist account of understanding in Sections 4.5-4.6 was the last part of this dissertation to be written, supplying a solution to a question that I had all but given up on satisfactorily resolving here. I thank the audience at the 4th annual conference on Scientific Understanding and Representation (SURe)—hosted at the Fordham University in April 2022—for helpful comments.

Despite this lengthy list of public appearances and collegial engagement, many parts of this thesis were written in what felt like a veritable intellectual vacuum. I do not rec-

ommend re-creating these vacuum-like conditions. They are not conducive for critical philosophy. For anyone in a similar situation, I suggest seeking much more help and guidance from peers and mentors. This project would have been the stronger for it.

Finally, to all those whom this dissertation disappoints, I apologize. I went the deepest I could go given the circumstances. At a certain point, one must choose to stop descending and begin the ascent. We are not paid to remain in the dark, even when we haven't found much light. I have come to think that the thing to do is to aim for a cleanly superficial job, to see if the simplest story worth telling is sufficiently satisfying. For even to tell this story, it will be complicated enough.

Preface

> So in the end, when one is doing philosophy, one gets to the point where one would like just to emit an inarticulate sound. — But such a sound is an expression only in a particular language-game, which now has to be described.
>
> –Wittgenstein, *Philosophical Investigations* 2009 [1953], §261

Philosophy provides a kind of therapy. For a long time, I found this picture of philosophy rather silly. Recently it has taken on a quite literal meaning for me. As the reader will soon discover, much of this thesis engages problems that do not find clear articulation in existing literature. These are problems that have arisen from my own pedagogical engagement with science and mathematics, going back to at least 2011. They are issues that have troubled me. Sometimes, it seems, we fashion our own fly bottles.

Owing to these idiosyncratic origins, much of the internal dialectic of this thesis consists in evaluating possible therapeutic responses to these problems. In many cases, I found no one in the literature defending the kinds of responses that seem *prima facie* viable to me. I was forced to concoct these positions out of whole cloth, motivating them wherever possible with remarks made by actual thinkers. Part of this is dissatisfying: it may appear as though I am tilting at windmills. Much of the thesis consists, as it were, in playing toy soldiers—something I'm afraid that, intellectually, I never quite grew out of. Nonetheless, it is the only way I could manage to alleviate the kinds of disturbances that have motivated this project. Where an interlocutor cannot be found, one must be invented in the imagination.

I doubt that my position will convince or appeal to anyone who does not already share a preponderance of my intellectual dispositions. This thesis is written primarily for them, in the hopes that it may alleviate the disturbances of a like-minded individual. To me, it has become evermore plausible that this is the most that philosophy achieves: it provides a particular kind of therapy for a particular kind of temperament. Different temperaments require different therapies. For those who find my position or predilections radically misguided, all I can hope is that they might be convinced of a conditional, namely that if they were disturbed in the particular ways that I am, then they would be inclined toward my position. At the very least, I report the remedies that I have found particularly therapeutic for the problems that I have found particularly disturbing. Like the pharmaceutical industry, I hope to find better remedies in the future.

Contemporary philosophy of science—especially of physics—has become so infected with metaphysics that it inherits the hopelessness of that task. For the empiricist-minded, the standard interpretive questions of philosophy of physics outstrip what we will ever have empirical access to decide. Philosophers of physics who pursue such questions are left with the methodology of metaphysics, hoping that differences in theoretical virtues can be rationally compelling. One goal of this thesis is to highlight a vast terrain of neglected questions about the cognitive significance of scientific and mathematical problem-solving. These questions evade metaphysical-infection, happily allowing even an empiricist to have a positive position. Metaphysics is at once the most powerful and most sterile of philosphical subdisciplines. It can summon anything into existence and thereby summons nothing.

Despite my focus on problem-solving, the framework of my arguments will still seem overly theoretical for some philosophers—particularly those sympathetic to pragmatism. Such philosophers might be skeptical of many emphases or distinctions I draw, including privileging non-practical over practical dimensions of the epistemic, knowledge-that over knowledge-how, and non-agentive aspects of understanding (especially my focus on computer states rather than actual humans). At a certain level of idealization and abstraction, I take these distinctions to be perfectly intelligible. Indeed, I take them to be parts of our intellectual practices, in both science and philosophy. We are clearly able to make these distinctions and have them be meaningful. I take it that these pragmatist philosophers are skeptical that such distinctions or emphases hold up at a more 'foundational' level. Such worries seem similar to the complaint that Quine leverages against the analytic–synthetic distinction. To me, it has always seemed that even if Quine is right that there is not really any analytic–synthetic distinction fundamentally, it is still an extremely useful and fruitful distinction within various idealized settings.

I find a rather frustrating hypocrisy in many philosophers' criticism of ideal theory in philosophy. These same critics are often extremely comfortable with idealizations made in the sciences. They extol these idealizations as paradigms of human rationality, often using them to motivate various grades of antirealism or to deflate more robust realisms. Yet, insofar as idealizations and toy models work well for the sciences, we have reason to think they will often work well in philosophy. It seems exceedingly two-faced to extol a particular problem-solving strategy when scientists do it and then scorn philosophers who follow suit. Hence, I proudly proclaim that I am engaged in a kind of ideal theory throughout this dissertation. One could try to de-idealize my account by adopting various non-ideal, foundational stances. These include (i) upholding a view-from-nowhere and appealing to categorical properties of agential rationality, perhaps privileging individuals or (ii) embodying my claims in actual agents, such as humans, perhaps making practices and communities foundational. I don't see any need to take a stand on these downstream issues here. Indeed, I think it is more befitting of the pragmatist-minded to not get overly wound up with how such foundational questions might shake out, at least in contexts where it seems unlikely to make much of a difference (which I believe is the case for the central questions in this dissertation). It is a strength of logic and computability theory that they abstract away from human agents, not a weakness.

Philosophers—across ideological spectra—are obsessed with trying to explain things. It is exceedingly common to frame the aim of a philosophical project as providing an explanation for some phenomena. I find such talk misleading: it presupposes that there is such a thing as *the explanation*, waiting out there in the world to be discovered. I do not take myself to be explaining anything here. I am describing aspects of our practices. By raising certain aspects to salience, I hope to alleviate some feelings of puzzlement. As Wittgenstein notes in an anthropological context, all we really need to do is describe carefully what is in front of us, and then the satisfaction that we thought could only come from an explanation comes of itself (1993 [1931], p. 120).

I have come to see that philosophy is an inherently normative enterprise. When philosophers say that they have given an explanation, they have given a description that they think ought to scratch a certain itch (either in themselves or others). Explanation is normatively-loaded description: it expresses an attitude of *being for being satisfied* by a particular answer. In general, when I defend a positive position, I express endorsement of a set of norms on what ought to satisfy you, dear reader. By rejecting some aspect of my positive position, you express endorsement of a different set of norms (at least in the interesting cases where we don't simply disagree about descriptive matters). Philosophical disagreement consists by and large in negotiating the norms governing what ought to satisfy us. For those who remain wedded to 'explanation' talk, such norms govern what counts as a philosophical explanation. In trying to make sense of the roles that various concepts play in human life, different philosophers disagree about the correct order or form of such explanations.

At the end of the day, every therapy has side-effects, along with ailments it just can't cure. We have to pick our poison, as it were, and different temperaments will choose differently. This thesis expresses the therapy that I have found so far to work for me. Along the way, it mentions a variety of other therapies that different readers might prefer, based on the nature of their own afflictions. Admittedly, I don't paint these therapies in a maximally welcoming light; they are not what I would recommend. I think there are good reasons against them, but the weight of my reasons might not seem reasonable to you. At a certain point, the reasons give out, and we are left with our temperaments.

And it is true, I wouldn't wish my temperament upon anyone. For that is to wish upon them an affliction. I'd rather them not worry about such matters; it would be better to go out and improve society. This is a work for the already-puzzled, who would like to be put back into unpuzzlement. If you are not already puzzled, I have a friendly suggestion: read no further. It is unlikely that you will find answers that satisfy you, at least not anytime soon. Barring the thrill of knowledge-seeking, it may be better to have never wondered.

One could perhaps try to tell an error theory as to why I ought not be troubled. But I will take my disturbances for granted. This is the method of intuitions. It is the germ from which philosophy sprouts, although not the soil that nourishes it. That soil is our practices. They are, as Wittgenstein notes, philosophy's raw materials (2009 [1953], §254).

A remark on content rather than form: I went into this project suspecting that there is something special about symmetry arguments in physics and chemistry. I came out the other end believing that symmetry arguments are a special case of what goes on in reformulations generally. The significant kind of reformulation changes what plans are available for problem-solving. Contrary to my expectations, the most important symmetry seems to be that which obtains between trivial notational variants. Trivial notational variants are related by a symmetry that preserves problem-solving plans. The important reformulations are the ones that break this symmetry. Symmetry arguments in physics and chemistry are epistemically significant because they provide an alternative path for problem-solving.

A remark on an aesthetic deficiency: Due to the practical necessity of publishing articles in academic philosophy, the chapters by-and-large can be read independently. In order to keep them relatively self-contained, I repeat some key definitions and aspects of my positive position, especially the notions of epistemic dependence relations (EDRs), intellectual significance, and conceptualism. I hope the reader will excuse the slight additional length this repetition has caused. A more elegant alternative might have been more burdensome to the reader.

On Reading More Widely

To a large extent, philosophy is about finding your tribe. We are searching for intellectual forebears and contemporaries who give voice to our own intuitions. Identifying these tribespeople can be much harder then it sounds, especially when your intuitions mesh with traditions that have been relatively suppressed in contemporary circles.

What has been particularly unsettling for me is the number of intellectual near-misses in my life. At Pittsburgh, I was not aware of the relevance of Brandom's work at all. At Cambridge, I did not encounter Huw Price. At Michigan, I did not read any Gibbard until more than five years in. While visiting NYU and USC in the Fall of 2019, I had yet to realize my deep sympathies for expressivism. Accordingly, I had no idea that Hartry Field was working on related issues; nor did I appreciate work by Horwich on truth and meaning; nor did I have any contact with Mark Schroeder. As such, much of the last year and a half has felt like playing a game of intellectual catch-up. What a world it would be if we had better 'relevance-detectors,' matching our intellectual temperaments to recommended reading. Of course, I bear responsibility for not doing due diligence. A life-long lesson if there ever was one: intellectual allies can be found in unsuspecting places!

Then again, perhaps one has to come close enough to these ideas on their own before they are sufficiently receptive to them. I recall sitting through a talk on expressivism just two-months into graduate school (the only such talk I've seen since, I'm afraid!). Although I thought parts of it might be relevant for addressing questions in philosophy of math, broader implications of the view did not occur to me. In the Fall of 2020, I even started reading a book by Price, only to set it down thinking the position too extreme in various ways. Yet at the same juncture, I was immediately gripped by norm-expressivism. At different stages, we seem receptive only to particular modes of presentation.

Inductively, my intellectual near-misses make me seriously concerned that the same sort of thing is happening right now: there presumably remain whole swaths of philosophy deeply connected to my own work and pattern of thought that I am simply unaware of. Given that I have been at this now in some capacity for 14 years, it is a rather humbling thought. I wonder how many other disciplines face a similar challenge of perpetual forgetfulness. Perhaps it is restricted to those disciplines where progress is not linear, where fads come and go. These experiences make me extremely grateful for the ideas that I *did* encounter at the right place and time, especially constructive empiricism, which I have held some version of since at least 2011.

My own methodological strategy is a cautionary tale. For many years, I myopically thought that philosophy of science held all the answers to the questions that puzzled me. I thought that all I had to do was read some cottage literatures on related issues while paying close attention to science. Before that, there was a period where I mainly just read science, thinking that philosophy could wait and wouldn't add much anyways. It is only recently that I have realized that what primarily puzzles me are aspects of normativity in science, and both philosophers of science and scientists are by-and-large wary of normativity. For instance, I have gotten much more out of reading Gibbard (2012) on meaning and normativity than papers in philosophy of science on theoretical equivalence. This is despite the fact that *prima facie*, the existing literature on theoretical equivalence would seem to be much more relevant for my project.

Recently (about a month before defending), I learned that Sosa (2015) defends many distinctions that I rely on in my own arguments. These include distinguishing between (i) non-practical vs. practical aspects of epistemic significance and (ii) constitutive vs. non-constitutive aims. Like me, Sosa characterizes the success conditions for an activity in terms of that activity's constitutive aims. This is yet another instance where reading outside philosophy of science would have helped resolve my philosophical puzzlement much sooner. Of course, there is something gratifying—and evidentially reassuring—about coming to these distinctions on your own, in light of independent examples and motivations.

While an undergraduate, I studied relatively few historical philosophers (matters have not improved much since!). Of those, Wittgenstein was the only one I read who I did not feel was systematically pulling the wool over my eyes. His disturbances struck me as honest grapplings with philosophical problems. His was the only philosophical methodology that I felt comfortable with. Yet, it was completely opaque to me who in the contemporary tradition had carried on Wittgenstein's methodology and background philosophical commitments. It is only recently that I have discovered one family of descendents in the Wittgensteinian tradition. Coming to understand how these family members fit within empiricism, pragmatism, and expressivism has been an incredibly illuminating project that will be ongoing for some time. Given the sense of having—yet again—merely scratched the surface of the philosophical terrain lying underneath, I cannot help but fret that much of what I say here will be *hopelessly naïve*. May it be steps in the ultimate direction that I would endorse once wiser from having studied more forebears in this grand tradition.

If I had to highlight just one misleading aspect of contemporary philosophy of science, it would be the unexamined dogma that we must look for the ontic truth-makers standing behind scientists' declarative sentences. Seeing my way out of this dogma was perhaps the most therapeutic intellectual development that I had in graduate school. Declarative sentences can perform more functional roles than representing reality. What a radical idea! This is a lesson that, in hindsight, I learned in 2013 from Tom Rickett's class on Wittgenstein, and then summarily forgot until 2020 (when I returned to the *Philosophical Investigations* to lift my pandemic-ridden spirits). My remembrance came alongside studying expressivism, which has cured me of many of my troubles.

Synopsis

By and large, philosophy of science has focused on *competing* theories and methodologies—cases where scientists are forced to choose between rival ontologies or inconsistent ways of understanding the world. Similarly, philosophy of mathematics has focused extensively on rival foundations for mathematics. Here, I show that compatible reformulations are no less important than competing theories or interpretations. Even if scientists were to reach a "final theory," they would still have multiple ways of representing and understanding the world. Likewise, it is exceedingly common in mathematics to have multiple ways of proving the same theorem or solving the same problem, often arising from quite disparate parts of mathematics. In both disciplines, it is essential to understand how reformulations contribute to intellectual progress.[2]

Chapter 1 begins by introducing questions about the nature and value of compatible formulations. Many reformulations seem to constitute a kind of intellectual advance. But how, and under what conditions? To answer these questions, I canvas a continuum of philosophical accounts of reformulation. These fall into four big 'isms': instrumentalism, conceptualism, explanationism, and fundamentalism.[3] The most deflationary position—instrumentalism—holds that some reformulations are merely instrumentally or practically valuable, in virtue of aiding problem-solving. At the other extreme, fundamentalism holds that some reformulations are metaphysically valuable in virtue of providing a more fundamental language for describing reality. Both conceptualism and explanationism provide middle ground positions between these. Explanationism contends that a reformulation is significant when it provides a better explanation. Finally, my preferred position—conceptualism—claims that a reformulation is significant when it provides an epistemically distinct plan for problem-solving. In virtue of how they structure problem-solving, some plans are epistemically better than others.

Section 1.5 introduces a variety of examples of reformulations from science and mathematics. Admittedly, I focus on cases of applied mathematics in science. This is largely for convenience: *mutatis mutandis*, it should be clear how my considerations generalize to many examples in pure mathematics. The contingent fact that these models are approximately instantiated in physical reality does not play a philosophically significant role in the dialectics I consider.

Chapter 2 ("Between Instrumentalism and Fundamentalism about Reformulation") defends three desiderata for a satisfying account of reformulations. Any such account should (i) distinguish trivial from significant reformulations in a way that (ii) applies to the local context of solving particular problems while (iii) using criteria that are epistemically accessible. I argue that instrumentalism fails to meet the first two desiderata, while fundamentalism fails to meet the third. I then show how conceptualism satisfies all three by distinguishing trivial and significant reformulations on the basis of epistemic structure.

[2]Hunt (2016) introduces a less general idea of 'compatible explanations' in applied mathematics.

[3]Hunt (2021a) explores part of this dialectical landscape in the context of an important class of symmetry arguments in physics and chemistry stemming from the Wigner–Eckart theorem for calculating matrix elements of quantum-mechanical operators.

Significant reformulations provide epistemically distinct problem-solving plans: they differ in what agents need to know to solve problems. In contrast, trivial reformulations differ only in notation, providing equivalent problem-solving plans. I briefly introduce explanationism as a rival middle ground position, noting some of its inherent disadvantages compared to conceptualism.

Chapter 3 ("Understanding and Equivalent Reformulations") provides a more detailed argument against explanationism by using cases of theoretical equivalence in physics, such as the Lagrangian and Hamiltonian formulations of classical mechanics. In these cases, two compatible formulations describe exactly the same states of affairs. Since explanations themselves ultimately refer to states of affairs, equivalent reformulations provide the same explanations. Alongside theoretically equivalent reformulations, I also consider diagrammatic reformulations, such as the use of Feynman diagrams in particle physics. These provide a further counterexample to explanationism.[4]

In Chapter 4 ("Reformulating through Symmetry"), I show that explanationism also fails to account for the intellectual significance of a common symmetry-based reformulation in quantum chemistry and solid-state physics. I argue that explanationism can accommodate the value of this reformulation only if it deploys substantial metaphysical commitments that my position avoids. At the same time, I provide a detailed analysis of how symmetry arguments improve scientific understanding. They do this principally through two mechanisms: modularization and unification (introduced in Section 1.4). Symmetry arguments modularize a problem into independently-treatable subproblems. Additionally, they unify different problems into symmetry-related classes. In both cases, this clarifies what we need to know to understand properties of the system of interest, thereby deepening our understanding.[5] The chapter closes with an expressivist account of what it means for one formulation to provide better understanding than another. When we judge that one formulation provides better understanding, we express an attitude of *being for intellectually preferring* that formulation.

Chapter 5 ("Reformulation as a Constitutive Aim of Science") further articulates the epistemic value of reformulating while providing a more detailed rebuttal of instrumentalism. I argue that science and mathematics constitutively aim at solving all possible

[4]Much of Chapter 3 appears in Hunt (2021b), excluding Section 3.7's case study on Feynman diagrams.
[5]These parts of Chapter 4 and Section 3.2 appear in Hunt (forthcoming).

problems within their respective domains. Ideally, a complete scientific theory would be able to solve all possible empirical problems. I refer to this constitutive aim as *planning adequacy*: it defines a minimal success criterion for science—and mathematics—in the futuristic ideal of limitless inquiry. Planning adequacy requires that agents can plan for any possible problem-solving context. In order to accomplish this aim, it is necessary to clarify what we need to know to solve problems. Clarifying epistemic structure requires reformulating. In this way, reformulating becomes a constitutive aim of science. Reformulations are not merely instrumentally valuable for this aim: they constitute its realization. Along the way, I distinguish between practical (or pragmatic) aspects of problem-solving and non-practical, epistemic aspects. These latter aspects are what I call *intellectual*.

Finally, Chapter 6 ("Making it Manifest: The Intellectual Value of Good Variables") uses a particularly compelling objection from fundamentalism as an opportunity to further develop conceptualism. Numerous examples from physics suggest that some reformulations are significant because they make fundamental properties manifest or perspicuous. Across the sciences, good variable choices lead to insights that bad variable choices obscure. Fundamentalism suggests that the value of making such properties manifest is that we thereby approach a more fundamental language for describing reality. This argument places pressure on a central desideratum from Chapter 2, namely a desire to avoid substantial metaphysical commitments in accounting for the value of reformulations.

I first provide an account of what it means to make a property manifest. In a given epistemic circumstance, a property is manifest provided that an agent in that circumstance ought to infer that the property obtains. This account naturally leads to characterization of what it means for an expression to "wear a property on the sleeves." An expression wears a property on the sleeves when a suitably informed agent ought to infer that property solely on the basis of other properties that the expression makes manifest.

When we make a property manifest, we simultaneously rule out more epistemically possible solutions than otherwise. This has the effect of structuring our search space for a problem. Insofar as we aim to solve problems, it is epistemically valuable to rule out as many epistemically possible solutions as possible (since doing so constitutes approaching a solution). We thereby arrive at a non-metaphysical account of the value of making properties manifest. Other things equal, it is intellectually valuable to make a property manifest because doing so rules out more epistemically possible solutions.

A fundamentalist might nevertheless hanker after a vindication of scientists' ordinary claims about one language being more fundamental than another. To satisfy this craving while respecting my own philosophical scruples, I develop a non-metaphysical, expressivist account of fundamentality. To judge that one formulation is more fundamental than another is to express an attitude of *being for privileging* the former formulation. Throughout the chapter, I illustrate and defend various facets of my positive position using numerous examples arising from natural language, graph theory, normal forms in logic, manifest Lorentz covariance, Cartesian vs. polar coordinates, gauge choices in quantum field theory, the hidden hyperspherical symmetry of the hydrogen atom, and hidden symmetries in $\mathscr{N} = 4$ super Yang–Mills theory.

Unlike instrumentalism, conceptualism makes sense of the intellectual significance of reformulations. Unlike explanationism and fundamentalism, conceptualism does so without embroiling us in controversial metaphysics. The account that emerges is one that places problem-solving at the heart of rational inquiry, emphasizing the functional roles played by scientific and mathematical concepts.

Abstract

Science and mathematics continually change in their tools, methods, and concepts. Many of these changes are not just modifications but progress—steps to be admired. But what constitutes progress? This dissertation addresses one central source of intellectual advancement in both disciplines: reformulating a problem-solving plan into a new, logically compatible one. For short, I call these cases of compatible problem-solving plans "reformulations."

Two aspects of reformulations are puzzling. First, reformulating is often unnecessary. Given that we could already solve a problem using an older formulation, what do we gain by reformulating? Second, some reformulations are genuinely trivial or insignificant. Merely replacing one symbol with another does not lead to intellectual progress. What distinguishes significant reformulations from trivial ones?

According to what I call *conceptualism* (or *conceptual empiricism*), reformulations are intellectually significant when they provide a different plan for solving problems. Significant reformulations provide inferentially different routes to the same solution. In contrast, trivial reformulations provide exactly the same problem-solving plans, and hence they do not change our understanding. This answers the second question about what distinguishes trivial from significant reformulations. However, the first question remains: what makes a new way of solving an old problem valuable?

Here, a bevy of practical considerations come to mind: one formulation might be faster, less complicated, or use more familiar concepts. According to *instrumentalism*, these practical benefits are all there is to reformulating. Some reformulations are simply more instrumentally valuable for meeting the aims of science than others. At another extreme, *fundamentalism* contends that a reformulation is valuable when it provides a more fundamental description of reality. According to this view, some reformulations directly contribute to the metaphysical aim of carving reality at its joints.

Conceptualism develops a middle ground between instrumentalism and fundamentalism, preserving their benefits without their costs. I argue that the epistemic value of significant reformulations does not reduce to either practical or metaphysical value. Reformulations are valuable because they are a constitutive part of problem-solving. Both science and mathematics aim at solving all possible problems within their respective domains. Meeting this aim requires being able to plan for any possible problem-solving context, and this requires reformulating. By reformulating, we clarify what we need to know to solve problems.

Still, one might wonder whether the value of reformulations requires underlying differences in explanatory power. According to *explanationism*, a reformulation is valuable only when it provides a better explanation. Explanationism stands as a rival middle ground position to my own. However, it faces numerous counterexamples. In many cases, two reformulations provide the *same explanation* while nonetheless providing different ways of understanding a phenomenon. Hence, reformulating can be valuable even when neither formulation is more explanatory.

Methodologically, I draw on a variety of case studies to support my account of reformulation. These range from classical mechanics to quantum chemistry, along with examples from mathematics. Symmetry arguments provide a paradigmatic example: the mathematics of symmetry groups radically recasts quantum mechanics and quantum chemistry. Nevertheless, elementary approaches exist that eschew this additional mathematical apparatus, solving problems in a more tedious but less mathematically-demanding manner. Further examples include reformulations of quantum field theory, Arabic vs. Roman numerals, and Fermat's little theorem in number theory. In each case, my account identifies how reformulations change and improve our understanding of science and mathematics.

Chapter 1:

Introduction

Throughout science, mathematics, and engineering, we often have multiple, compatible methods for solving problems. For each theory that scientists develop, they typically develop multiple ways of expressing or formulating its physical content. Often, the motivations for reformulating are practical: scientists wish to solve problems more quickly, simply, or elegantly. Sometimes, the aim is explicitly to clarify conceptual foundations, often by applying new mathematical techniques. Either way, the results of reformulating are a significant aspect of scientific progress. Reformulations often change how we understand the world, spawning new areas of research that probe the properties and scope of the reformulated theory. Similar remarks apply to reformulations in mathematics.

When it comes to scientific reformulations, physics supplies a wellspring of examples. Within classical mechanics, there are no less than five ways of formulating a large variety of problems. These include the Newtonian, Hamiltonian, Lagrangian, Hamilton–Jacobi, and Routhian formulations of classical mechanics (Abraham and Marsden 1978; Arnold 1989). They differ in their mathematical strategies for solving the equations of motion for classical systems, and—within their shared domain of applicability—they describe the same physical states of affairs. Similarly, nonrelativistic quantum mechanics can be formulated in a variety of distinct mathematical garb, including wave mechanics, matrix mechanics, density operators, and path integrals (Styer et al. 2002). Other examples include the fiber bundle formulation of classical and quantum field theories (Healey 2007), the use of topology to characterize broken symmetry in condensed matter physics (Sethna 2006, Ch. 9), applications of de Rham cohomology in general relativity (Torre 1997), and various formulations of thermodynamics in terms of different ensembles (including the Helmholtz, Gibbs, and grand free energies).

In each of these cases, the various formulations are *compatible*: we are not forced to choose between them. Rather, we can accept and use them all. They do not disagree about the way the world is. They do not posit competing ontologies, nor offer competing descriptions or predictions. Instead, they provide logically consistent problem-solving procedures for a shared class of problems. For short, I will refer to such compatible formulations as *reformulations*.

Clearly, reformulations like these provide different ways of looking at their shared subject matter. I take it as a datum of intellectual life that some of these reformulations constitute progress. Nevertheless, it is difficult to characterize how reformulations lead to differences in understanding. This dissertation provides a positive account of how reformulations change and improve our understanding. In this chapter, Section 1.1 characterizes a philosophical puzzle that reformulations inspire. Next, Section 1.3 introduces a variety of philosophical responses, on a continuum from maximally deflationary to maximally metaphysically-committal. The account I develop occupies a middle ground between these extremes. It focuses on how reformulations improve our epistemic position with regards to solving problems, discussed in Section 1.4. Finally, Section 1.5 provides a variety of examples from science and mathematics that motivate and illustrate my account. Later chapters simultaneously elaborate my position while defending it from rivals.

1.1 A Puzzle about Compatible Reformulations

To date, much philosophy of science and metaphysics has focused on *competing* theories or formulations. In the case of competing theories or formulations, two approaches to the same problem are incompatible: at most one could be true. Typically, competing theories compete because they posit incompatible ontologies. The chief philosophical task becomes one of weighing the reasons for and against these incompatible ontologies. Bohmian vs. Everettian interpretations of quantum mechanics provide a paradigmatic example of this kind of philosophical problem. Since competing formulations posit rival ontologies, they straightforwardly lead to different ways of understanding the world. Thanks to these ontological differences, competing formulations do not generate the philosophical puzzle that interests me here.

In contrast, reformulations arise whenever we have two or more compatible approaches to the same theory or set of problems. The compatibility of the approaches means that we are not forced to choose between them. Instead, we can endorse them all, solving problems with whichever formulation we prefer. Unlike competing theories or methodologies, reformulations disagree about neither the way the world is nor the nature of mathematical facts.

This lack of ontological disagreement leads to a host of novel philosophical questions: what do we gain by recasting a theory or problem-solving procedure in new terms? How is it that reformulating changes our understanding? Why do we often care about or value reformulations? Should we care about them in the way we do? These kinds of philosophical questions remain underexplored in epistemology, metaphysics, and the philosophy of science. In the first instance, they are not about ontology or the nature of reality. Instead, they concern how scientific and mathematical agents make intellectual progress by reformulating their problem-solving procedures.

Reformulations are puzzling for at least the following reason. For a given problem, no particular formulation is necessary for providing a solution. Any compatible formulation would suffice. In this way, each compatible formulation seems to render the others dispensable, at least for the purposes of problem-solving. Nevertheless, many reformulations seem to constitute a particular kind of intellectual progress. They characteristically deepen our understanding.

Such changes in understanding display at least three interrelated features: i) they are objective, ii) they can be meaningfully distinguished from practical (or 'pragmatic') concerns, and iii) they constitute changes to our epistemic position. Such changes are what I will call *intellectually significant*: they concern objective, non-practical, epistemic dimensions of understanding.[1] The primary philosophical challenge of this dissertation is to characterize and illuminate how some reformulations generate *intellectual differences*, i.e. objective, non-practical, and epistemic differences in understanding. Section 2.3 proposes three desiderata that any satisfying account of reformulations should satisfy. Having isolated some characteristic intellectual differences, we can then consider how some of them constitute intellectual improvements: they not only change our understanding but improve it.

Generally, I will not provide an argument for interpreting any particular example as

[1]For a similar use of 'intellectual,' see Sosa (2015, p. 45).

being a case of compatible reformulations. In all the examples I consider, it will be *prima facie* plausible that they are compatible reformulations rather than competing theories. Indeed, cases of mathematical and diagrammatic reformulations are generally *best* interpreted as being compatible. In mathematical reformulations, we change the mathematics that we apply and this has no implications for physical ontology. Moreover, I take it that different parts of mathematics do not compete with each other (Hunt 2016, p. 462). In diagrammatic reformulations, we similarly change the representational tools we are using, and again this does not alter the ontology expressed by the theory. Nevertheless, some cases will be more controversial. For instance, some may be tempted to view Newtonian and Lagrangian mechanics as providing competing ontologies for classical mechanics (e.g. a force-based vs. potential-based ontology, respectively). Under this interpretation, the puzzle of compatible reformulations does not arise, and the philosophical task falls squarely within traditional debates concerning which ontology we ought to believe more likely to be true. How to approach competing formulations or interpretations is already well-catered for by existing literatures. My focus here is on what we should say in the event that two problem-solving procedures are genuinely compatible.

1.2 Intellectual Significance

Why introduce this jargon word, the 'intellectual'? In short, it functions as a convenient shorthand, delimiting the scope of my investigation. I am interested in objective aspects of understanding that are non-practical and epistemic. Unlike 'explanation,' the English noun 'understanding' has neither an adjectival form parallel to 'explanatory' nor an adverbial form parallel to 'explanatorily.' As my focus is on differences in understanding, it seems grammatically convenient to call these 'intellectual differences.' We can then denote differences in understanding that—for my purposes—are philosophically significant as being 'intellectually significant differences.' Throughout, one must keep in mind that I do not mean just any difference in understanding. Many such differences are practical in nature. The *intellectual* picks out a particular facet of scientific and mathematical understanding.

Why not then just stick with 'epistemic' and its kin? Indeed, what I am calling 'intellectual' includes what many would call 'purely epistemic' matters (Sosa 2015, p. 172).

Unfortunately, the word 'epistemic' is highly contested (S. Cohen 2016). Many think that epistemic matters naturally include issues that I classify as practical, such as problem-solving speed or efficiency (Goldman 1986, p. 122). If one wants to classify such practical matters as being epistemic, I am inclined to let them, rather than to quibble over terminology. Nonetheless, I contend that we can isolate a non-practical dimension of the epistemic. My usage of 'intellectual' builds in this isolation.[2] If I were to look hard enough, I might find that 'intellectual' has also already become highly contested. Nevertheless, it seems much less common in the literatures I engage with here, lending itself more readily to a stipulative characterization. As we'll soon see, I flirted with using 'conceptual' instead. Here the problem is not so much that the word is contested but that it applies so broadly as to lose clear sense.

It is difficult to non-contentiously define the three interrelated adjectives that characterize the intellectual, namely 'objective,' 'non-practical,' and 'epistemic.' Each is to be contrasted with another controversial or unwieldy notion, namely 'subjective,' 'practical' (or 'pragmatic'), and 'non-epistemic' (including moral and aesthetic). The contrasts I intend will become clear as we proceed through examples. If the reader can't wait for a better grip on these contrasts, hopefully the following suffices for now: matters of taste are not objective. Formal, syntactic features of theory formulations or problem-solving plans are objective. Epistemic matters concern justification and knowledge (a kind of true belief).[3] Practical matters concern features of agents that in-principle do not make a difference for acquiring knowledge. These include matters of convenience, speed, efficiency, and effort.

In distinguishing the non-practical from the practical, I intend to remain neutral on

[2]Sections 2.6.3 and 5.6-5.7 contrast non-practical and practical dimensions of the epistemic.

[3]What is knowledge, you might ask? Well, you know it when you see it! More seriously, the arguments in this dissertation are compatible with a wide variety of different accounts of knowledge. If pressed for my own view of the matter, I am inclined to agree with Wittgenstein (2009 [1949], §307, 310): when an agent judges that they know P, they indicate that they currently do not doubt P. Dressed up in more contemporary garb, they express an attitude of *being for not doubting P*. This attitude is very different from expressing that there is no doubt that could possibly be had. Of course, whether you know P depends on whether P is true.[4]

[4]Well then, what is truth, you might ask? That's also highly controversial. In some contexts (dealing with observables), I will implicitly take 'truth' to mean correspondence to states of affairs in reality. In other contexts (especially when values and norms are in play), 'truth' will mean simply that we can apply a disquotation schema: the sentence 'P' is true if and only if P. If you vindicate asserting 'P,' you vindicate asserting that 'P is true.' Such processes of vindication do not count as objective in the sense of correspondence, but still in the sense that they are not matters of taste. These judgments are governed by higher-order norms such that we do not take them to depend on features of us. Field (2018, p. 16) calls this a kind of 'counterfactual objectivity.'

whether philosophers ought to ultimately privilege practical matters. For instance, maybe all intellectual aspects of agentive activity are ultimately socially grounded in some deeper sense. Either way, we can distinguish between practical and non-practical matters within some background practices such as science or mathematics. This distinction is compatible with taking these background practices themselves as fundamental for philosophical purposes. To do so would be to adopt a form of pragmatism, which Brandom defines as "a generic expression that picks out a family of views asserting various senses in which practice and the practical may be taken to deserve explanatory pride of place" (2011, p. 58).

1.3 A Continuum of Solutions

To resolve this puzzle about compatible reformulations, I will consider a number of putative solutions. We can visualize these solutions as lying along a continuum from maximally deflationary to maximally inflationary, i.e. involving the adoption of substantial metaphysical commitments. My goal will be to defend a middle ground position between these extremes, which I will call *conceptualism*.[5] It focuses on the way in which significant reformulations provide new concepts that genuinely alter our problem-solving structures. As with most middle grounds positions, conceptualism must steer a path between the Scylla of deflationary solutions and the Charybdis metaphysically inflationary ones. As the issues come into focus, a competing middle ground position—explanationism—will emerge as my strongest opponent. Chapters 3 and 4 defend conceptualism against explanationism.

The most deflationary solution is to deny that there is any puzzle at all. Perhaps when it comes to reformulations, the only differences that arise are ones of mere convenience. According to what I will call *conventionalism*, this is all there is to say about reformulations. Reformulations provide convenient footholds for forging ahead, facilitating the solution of problems we could solve with other methods if only we were willing to sacrifice the time and energy. Conventionalism holds that there is nothing deep or intellectually

[5] I call this view "conceptualism" to emphasize the role that concepts play in theory reformulation and understanding. By "concepts," I include what Kenneth Manders (2008b, unpublished) calls "expressive means." These include the mathematical, linguistic, diagrammatic, and notational resources we use to express theories. More precisely, I endorse Gibbard's (2012, Ch. 2) distinction between properties and concepts, where concepts are the content of thoughts. Despite some interesting analogies, I do not intend to endorse scholastic or early modern *conceptualism about universals*.

significant about reformulations. They merely amount to a different choice of convention, no different in kind than a change in notation. Seen through this lens, the seeming intellectual triumphs of wholesale theoretical reformulations are simply one notational change after another, convenience piled atop convenience. Conventionalism holds that reformulations are different only in degree—rather than kind—from trivial notational changes (Section 2.4.1 characterizes conventionalism in more detail).

I doubt that any philosopher or scientist would defend conventionalism in all cases of reformulations. For conventionalism faces a serious problem. It fails to save the intellectual phenomena staring us in the face. As the history and practice of science and mathematics shows, we care about reformulations an awful lot. Section 1.5 provides many examples. It is no doubt true that partly why we care so much about reformulations is due to the convenience that they often provide. The conventionalist has no problem saving this aspect of the story. However, focusing on convenience alone results in an incredibly bleak picture of our intellectual enterprise. No reformulation would count as genuinely increasing (or even changing!) our understanding of a given theory. It would be mere convenience all the way down, invalidating our usual attributions of depth or insight to paradigmatic reformulations. Conventionalism leaves us with a bland and disappointing error theory about the value of reformulations.

Section 2.4 considers a more sophisticated deflationary position, focusing on the general instrumental value of reformulations. I will call this position *instrumentalism*. It contends that the value of reformulations reduces entirely to their instrumental value for the basic aims of science (or mathematics). Chapter 5 provides an extended argument against instrumentalism. I show that reformulation itself constitutes an aim of science. Reformulation therefore acquires final value and is hence not reducible to the other basic aims of science.

At the other extreme of our continuum lies a metaphysical picture like that of David Lewis's. Lewis posits that some properties belong to an elite set of *perfectly natural properties*, with physics aiming to provide a partial inventory of these properties (1983, pp. 357, 364). Ted Sider speaks instead of a theory's conceptual structure, which must match the structure of reality in order for the theory to be "fully successful" (2011, p. vii). Sider's framework suggests that two formulations of a theory can state the same truths about the world while nonetheless disagreeing about which concepts are more fundamental, i.e. more joint-carving (2011, p. 5). These pictures motivate a metaphysically-committal solution to the problem of compatible formulations. According to what I will call *fundamentalism* (Section 2.5), a compatible formulation does better the closer it comes to a canonical language that carves nature at its joints. Chapters 2 and 6 argue that the metaphysical commitments of fundamentalism are not necessary. We can provide a satisfying account of the intellectual significance of reformulations without the fundamentalist's ontological commitments.

A less metaphysically expansive solution appeals to putative explanatory differences between reformulations. According to what I will call *explanationism*, reformulations are intellectually valuable when they provide alternative explanations. Due to the vast number of different accounts of scientific explanation, explanationism provides a schema, to be filled in with a particular account of explanation. Different accounts of explanation give rise to different versions of explanationism.[6] For this reason, it is logically difficult to argue decisively against explanationism. Chapter 3 provides a general argument against a wide class of explanationist accounts. I consider a special case of compatible formulations, namely formulations that are theoretically equivalent. I argue that theoretically equivalent formulations provide *the same* explanation but nevertheless can have intellectually significant differences. Chapter 4 broadens the scope of this argument to a wider class of compatible formulations. Doing so requires engaging with particular accounts of explanation. I rebut four general families of explanationist accounts, based on different notions of explanatory relevance.

The core of my argument against explanationism is the following: reformulations seem to manifest a number of differences that *prima facie* do not appear to be matters of explanation. Instead, these differences involve changes to the epistemic structure of problem-solving. They involve changes to how scientists and mathematicians go about structuring a search space or organizing information. Hence, I believe that a general account of reformulations requires focusing on how formulations structure problem-solving. Answering explanatory why-questions is, after all, just one kind of problem.

I call my proposed middle ground solution *conceptualism* (or more suggestively, "conceptual empiricism").[7] I aim to characterize the intellectual significance of reformulations

[6] Sections 2.7 and 3.2 characterize explanationism in more detail.

[7] Broadly, my position combines elements from van Fraassen's (1980) constructive empiricism and what Brandom calls "conceptual pragmatism" (2001, p. 4). Regrettably, I do not spell out these connections in

by appealing to their *expressive means*, i.e. their logical, mathematical, diagrammatic, and linguistic resources.[8] I will argue that reformulations are significant when they introduce concepts that restructure problem-solving. What matters are the different functional roles that different formulations support. By focusing on concepts rather than ontology, conceptualism stays within the meager ontological commitments of austere empiricist conceptions of science and mathematics. Sections 2.6, 3.5, and 6.4 characterize different facets of conceptualism in more detail.

Of course, there is nothing to prevent either a fundamentalist or an explanationist from adopting my conceptualist analysis and simply wanting to add more. A fundamentalist might wish to append additional commitments to fundamental structure. An explanationist might wish to append additional commitments to explanatory differences. My view is not incompatible with either of these augmentation strategies. Instead, conceptualism is simply incompatible with either fundamentalism or explanationism *being the end of the story* when it comes to the intellectual value of reformulations. My goal is to show that on their own, various versions of conventionalism, fundamentalism, and explanationism provide inadequate accounts of reformulation. These negative arguments motivate a need for conceptualism as a positive account of the value of reformulations.

1.4 Characteristic Differences in Epistemic Structure

In Chapter 2, I will argue that differences in problem-solving structure amount to differences in what we *need to know* or what *suffices to know* to solve a problem. I will refer to facts about what we need to know or what suffices to know as *epistemic dependence relations* (EDRs).[9] The *relata* of these relations are the inputs and output(s) of a reasoning process or problem-solving procedure. Schematically, EDRs provide answers to the following kinds of questions: do I *need* to know B in order to know C? Does knowledge of D *suffice* for knowing E?

As a simple illustration of epistemic dependence relations, consider a toy example

detail, although I hope they become clear for the cognoscenti.

[8] I borrow this expression from Manders (2008a), who introduces it in the context of diagrammatic reasoning in Euclidean geometry but does not provide a definition.

[9] EDRs also encompass facts about what we *don't* need to know and what does *not* suffice to know. Some may prefer to call these kinds of relations "epistemic independence relations."

from arithmetic: calculating the absolute value of the product of two integers, $|xy|$.[10] One formulation of this problem involves first calculating x times y and then taking the absolute value. This involves knowing the signs of both integers. Alternatively, one could reformulate this problem by recognizing that the absolute value of a product equals the product of the absolute values: $|xy| = |x|\,|y|$. In this formulation, we don't need to know the sign of each integer. It suffices to know their absolute values. Hence, this reformulation uses different EDRs to solve the problem. Conceptualism claims that in cases like this, we not only solve the problem differently, but also gain a different understanding of the solution.

Although epistemic dependence relations are *sui generis*, conceptualism seeks to classify them into different families, based on their functional roles. I will refer to these as different kinds of EDRs, distinguished by different intellectually significant properties. These properties characterize different aspects of problem-solving structure. In the case studies I examine from physics, chemistry, and mathematics, three kinds of EDRs play starring roles: modularization, unification, and uniformity of treatment. I consider these in turn.

Some epistemic dependence relations perform the functional role of *modularizing*: they show how to break a problem into separately treatable (or at least partially decoupled) sub-problems. Modularization occurs when a formulation solves a problem by decomposing it into separately treatable sub-problems. In the context of answering explanatory why-questions, this involves decomposing a why-question into separately treatable why-questions. Common examples of modularization include recursion relations and separation of degrees of freedom. Modularization arises outside the natural sciences as well, including modular programming in computer science (Avigad 2015). In Chapter 4, I argue that modularization lies outside the scope of philosophical theories of scientific explanation. It thereby illustrates an intellectually significant difference that is not an explanatory difference.

Other EDRs perform the functional role of *unifying*: they show how the same answer or result applies to a range of different problems or phenomena. For instance, a symmetry-based derivation can unify the energy spectra of two or more molecules that have the same symmetry group. Unification is a kind of EDR because it provides information that is

[10] I thank Dave Baker for suggesting this example.

sufficient for solving problems about an entire class of particular systems. This shows that problems that *prima facie* seemed to require separate answers can actually be answered collectively. In general, unification allows a sub-problem to be solved once and then used in subsequent applications. Like modularization, unification describes a class of EDRs, since many different EDRs play the functional role of unifying.

Unlike modularization, unification admits a natural interpretation in terms of explanations. Successful unifications often answer why-questions such as *why do all of these systems display this behavior?*[11] However, many formulations that fail to unify nevertheless succeed at solving the relevant class of problems case-by-case for each system of interest (often through a uniform, but not unified, treatment). We will see this again and again with the elementary approaches to various problems considered below. On some accounts of explanation, the resulting disunified conjunction of explanations—one for each system—is a fine way to answer an overarching why-question.

In Chapter 4, I argue that even if this conjunction of answers is explanatory, there is a crucial difference between a unified answer and a disunified, conjunctive answer. An account of reformulations must do this difference justice. Moreover, it remains controversial whether this aspect of unification is genuinely an aspect of explanation. For my purposes, settling this controversy does not matter. If unification is explanatory, so much the better for accounts of explanation that accommodate it. If it is non-explanatory, it nevertheless fits into the account of scientific understanding that I provide.

Finally, *uniformity of treatment* characterizes how a single problem-solving plan can apply to different kinds of problems, treating these problems uniformly. It is an epistemic dependence relation because it shows that applying a certain set of steps is sufficient for solving these problem-types. Unlike in cases of unification, uniformly-treated systems can still display different behavior and therefore involve different answers or wholly different problems. In an instance of unavoidable terminological infelicity, what I am calling "uniformity of treatment" is an aspect of Kitcher's (1989) account of *unification*. In those cases where one Kitcherian argument pattern applies to many different kinds of systems, we have what I would call a uniform treatment. In such cases, a single argument pattern is instantiated differently for the different kinds of systems. This distinguishes uniformity

[11] Indeed, what I call "unification" is similar to what Bob Batterman calls answers to type II why questions, which ask why a certain pattern generally obtains across different systems (2002, p. 23). It is also similar to what Alan Baker calls "topic generality" (2017, p. 200).

of treatment from what I am calling "unification." In cases of unification, we have a single instantiation of an argument pattern that applies to a variety of different systems.

For instance, group representation theory provides a uniform treatment of the spectroscopic properties of elementary particles, nuclei, atoms, and molecules based on their symmetries (these symmetries provide Casimir operators whose eigenvalues can be used to label states). Systems with different symmetries require different accounts of these properties, but the group theoretic solution procedure applies uniformly. However, when it comes to figuring out how energy levels are ordered from least energetic to most energetic, a fragment of the group theoretic approach does *not* provide a uniform treatment. Instead, without solving for the eigenvalues characterizing these levels, the best that symmetry alone can do is provide *ad hoc* geometric arguments based on the relative spatial overlap of orbital density functions and constituent atoms or molecules. In contrast, for some problem-types there is a non-group theoretic symmetry-based approach that uniformly treats the ordering of the energy-levels as a function of symmetry.

As another example of a failure of uniformity, consider *Ansatz*'s, used throughout physics to solve problems. An Ansatz is an educated guess at the form of a solution, which once guessed can be verified to hold true. Although some Ansatz's (such as separation of variables in differential equations) are applicable in a wide range of circumstances, it is typically possible that an Ansatz might fail. In this way, formulations that use Ansatz's often do not treat a class of problems uniformly, even if they are nevertheless able to solve each problem within a class.

Conceptualism treats each of these three features—modularization, unification, and uniformity of treatment—as intellectually significant aspects of reformulations. Although they typically make solving problems and providing explanations more convenient, they have a non-practical, epistemic aspect that goes beyond mere convenience. These intellectually significant aspects arise from illuminating or manifesting an alternative epistemic structure. They are valuable at least because they are epistemically valuable. A fundamentalist will wish to go further and base these features on objective joints in nature, while conceptualism remains agnostic about this further commitment.

Whereas convenience or instrumental value is an extrinsic feature of formulations, intellectually significant features are intrinsic. Discerning intellectually significant differences involves comparing at least two different formulations, but the epistemic dependence relations themselves are non-comparative. They are an intrinsic feature of that formulation's expressive means—objective features of its epistemic structure. Whether or not a formulation modularizes, unifies, or provides a uniform treatment of a problem is settled by that formulation alone, in isolation from considering other formulations. In contrast, assessments of convenience are inherently comparative: no formulation is convenient *simpliciter*, it is only either more-or-less convenient than another formulation. Assessing how convenient an approach is requires some external metric, such as a clock or counting procedure.

For instance, solving a problem on a faster computer is convenient because it is faster than solving it on a slower computer.[12] In many of the cases discussed in Section 1.5, we can use the number of matrix elements that a formulation must compute as a measure of its convenience. Reformulations that reduce the number of matrix elements (or replace matrix element calculations with simple arithmetic) can be viewed as more convenient.[13] In sum, intellectually significant features must be intrinsic aspects of a formulation, and differences between these intrinsic features account for intellectually significant differences between reformulations.

1.5 Reformulations in Science and Mathematics

Readers who are already convinced of the significance of compatible reformulations can feel free to skip or skim this section, moving to Chapter 2. There, Section 2.2 provides two simple examples of reformulations that suffice for articulating the relevant philosophical dialectic.

A wide variety of examples from science and mathematics motivates my account of reformulations. In physics and chemistry, reformulations range from classical mechanics to quantum field theory. Since my account applies uniformly to both reformulations in

[12]Note that this does not make knowledge of computational complexity a matter of mere convenience. For instance, it is intellectually significant to show that a problem can be solved in polynomial time. This provides an epistemic dependence relation, e.g. *this algorithm suffices to solve the problem in polynomial time* or, *we now know that this problem can be solved in polynomial time*. Presumably, this is more convenient than solving the problem in exponential time, but it is not merely more convenient: figuring out how to solve the problem more quickly is intellectually significant. For the same reason, knowing how to build a quicker computer is intellectually significant, even though using it to speed up a program is not. Section 5.7 considers these aspects of speed vs. intellectual significance in detail.

[13]In Chapter 4, we will see that there are reasons to question this simple metric for convenience. More sophisticated reformulations generally require much more complicated expressive means, which in some contexts mitigate the computational convenience that they provide. Chapter 2 will argue that assessments of convenience are agent-relative, while epistemic dependence relations are not.

science and mathematics, I sketch a few examples from pure mathematics as well. Later chapters discuss many of these examples in detail, using them to illustrate and defend conceptualism.

The puzzle about reformulations arises in each example below, and we will be interested in comparing and contrasting how conceptualism and its rivals treat these cases. Although the gains in convenience are often dramatic, the intellectual richness of these examples places pressure on any purely conventionalist response. Since these examples arise naturally in the practice of science and mathematics, they demonstrate that the puzzle about compatible reformulations is not an artifact of contrived examples. Many cases considered below involve reformulating a problem using symmetry, and I will argue that these are a special case of reformulations more generally.

To start with a well-known example, consider some of the intellectual differences between Newtonian and Lagrangian formulations of classical mechanics. For a large class of physical phenomena, both frameworks provide equally legitimate explanations. They use the same natural laws and the same initial conditions to derive a given *explanandum* (e.g. the position of a simple pendulum at a given time). In other words, both are committed to the same physical ontology and do not compete with each other; neither is closer to the truth than the other (although a fundamentalist might still wonder which one, if either, best describes the structure of a classical world). Nevertheless, these approaches provide different epistemic dependence relations. The Lagrangian approach provides three that are typically taken to be striking advantages: i) coordinate independence of the Euler–Lagrange equations; ii) elimination of constraint forces from many calculations; and iii) a uniform method for treating physical laws via an action principle (Goldstein et al. 2002, pp. 24–25, 35–36). These advantages are arguably responsible for physicists' strong preference for the Lagrangian formulation over the Newtonian. They make solving many problems much more convenient.

A conventionalist would say this is the end of the story: the Lagrangian reformulation matters merely because it is more convenient. A fundamentalist might argue that the Lagrangian approach carves nature more closely at its joints. Conceptualism acknowledges this is possible, but aims to account for the intellectual significance of Lagrangian mechanics independently of this further commitment. We can do so by interpreting each of the three advantages above as an epistemic dependence relation. The use of an action principle is an example of providing a uniform treatment, enabling systems as diverse as the motion of one free particle, an Atwood's machine, and a bead moving on a rotating wire to be treated under the same procedure.[14] The coordinate independence of the Euler–Lagrange equations enables us to work in generalized coordinates which contribute to this uniform treatment. Finally, eliminating constraint forces from the equations of motion amounts to modularization: we learn that we can separate the explanandum of interest (solving the equations of motion) from figuring out the constraint forces acting on the system. If we want to, we can still solve for the constraint forces as a separate step, but unlike in the Newtonian procedure, we do not need to know these forces to find the equations of motion.

Reformulations in Quantum Mechanics

Turning to a second example, consider one of the most prevalent problems in quantum physics and quantum chemistry: the calculation of matrix elements for various physical operators. These matrix elements represent a system's physical properties, including the likelihood of a transition from one state to the next. Using an *elementary approach*, each matrix element can be calculated via a corresponding inner product in Hilbert space, representing the states and operators of interest in some basis, e.g. position space, momentum space, etc. However, this elementary approach quickly becomes inconvenient, due to the sheer quantity of matrix elements. For instance, calculating the likelihood of a transition between a d-orbital and p-orbital under electric dipole radiation (i.e. first-order in perturbation theory) would naïvely involve calculating 45 matrix elements. To avoid this, a great deal of energy has gone into making these calculations as convenient as possible. This involves exploiting matrix element theorems such as selection rules (which characterize which matrix elements must vanish, as entailed by symmetries and conservation laws) and the Wigner–Eckart theorem, which separates out rotational degrees of freedom from other degrees of freedom. The rotational degrees of freedom can be cataloged separately in terms of Clebsch–Gordan coefficients. The other degrees of freedom are bundled into a *reduced matrix element*, which represents the particular physical de-

[14]Here is a summary of this uniform treatment: "a straightforward routine procedure can now be established for all problems of mechanics to which the Lagrangian formulation is applicable. We have only to write T and V in generalized coordinates, form L from them, and substitute in [the Euler–Lagrange equations] to obtain the equations of motion" (Goldstein et al. 2002, p. 24).

tails of the system at hand. For instance, the selection rules tell us that only nine of the 45 matrix elements do not necessarily equal zero. The Wigner–Eckart theorem then lets us calculate only one matrix element using the elementary approach, determining the rest by the rotational degrees of freedom.[15]

It is of course possible to interpret this reformulation of the matrix element problem as *merely convenient*. Indeed, striving for convenience provides much of the motivation for developing these techniques. However, it woefully mischaracterizes the knowledge that selection rules and the Wigner–Eckart theorem provide. The Wigner–Eckart theorem both modularizes and unifies the matrix element problem. It modularizes the matrix element problem by separating it into the following sub-problems: compute the Clebsch–Gordan coefficients for the relevant symmetry group (e.g. $SO(3)$), calculate one matrix element using an elementary approach, determine the reduced matrix element, and then apply the Clebsch–Gordan coefficients to determine all of the other matrix elements in this symmetry-related family. Matrix elements that share the same reduced matrix element are unified into a symmetry-based family, differing only in their rotational degrees of freedom. Even if we don't commit ourselves to joint-carving, these epistemic dependence relations are intuitively intellectually significant. They radically change what it suffices to know to solve for the family of 45 matrix elements.

Calculating matrix elements is also frequently aided by using symmetries to diagonalize matrices, especially in the context of external symmetry breaking. Even though matrices can be diagonalized numerically, these methods can slow down considerably as the matrices grow larger for more complex problems. In general, the computational time is proportional to the matrix-dimension cubed (McIntosh 1971, p. 78). By exploiting a physical system's symmetries, we can at least block-diagonalize a perturbation matrix, typically resulting in huge swaths of vanishing matrix elements. This modularizes the initial perturbation theory problem into separately treatable secular equations, one for each sub-block (Cornwell 1984, pp. 132–133).

Chapter 4 considers an extended case study from crystal field theory, which models the electronic properties of transition metal complexes. Surrounding a metal ion with ligands breaks the degeneracy of its valance electrons, causing previously degenerate or-

[15] For an extended discussion of the Wigner–Eckart theorem, analyzed in terms of conceptualism, see Hunt (2021a).

bitals to *split* into new energy levels with new degeneracies. For instance, when placed in an octahedral field, the five-fold degenerate d-orbitals of a nickel ion split into a more energetic two-fold degenerate level and a less energetic three-fold degenerate level. Naïvely solving for the eigenvalues would involve calculating 25 matrix elements. By taking advantage of the nickel ion's spherical symmetry and the crystal field's octahedral symmetry, we can construct symmetry-adapted basis functions for this calculation. In this symmetry-adapted basis, we modularize the original eigenvalue problem into two separate eigenvalue problems, one for each of two degenerate eigenspaces. We then have to calculate only two matrix elements, one for each eigenspace.

The group theoretic reformulation of external symmetry breaking provides other insights besides just the diagonalization of matrices. Representation theory describes how initially degenerate levels rearrange when placed within an environment of lower symmetry. It reformulates this problem in terms of the decomposition of representations of the initial symmetry group into a direct sum of irreducible representations of the lower symmetry group. Thus, it shows us how to solve for the number and degeneracy of the resulting energy levels, all without even solving for the eigenvalues. This provides another illustration of modularization: solving for the splitting and degeneracy of the levels can be treated separately from solving for the eigenvalues. It also illustrates unification: based on their symmetries, different systems can be organized into symmetry-based equivalence classes that display the same pattern of behavior. Just as a group's Clebsch–Gordan coefficients can be tabulated, these relationships between (i) the initial and final symmetry groups and (ii) the energy levels can be calculated once and tabulated (Bethe 1929, p. 143; Cotton 1990, pp. 264–265). Again, although these results can be incredibly convenient for calculations, it would seem to be completely inadequate to characterize them as merely convenient. Chapter 4 considers this group theoretic reformulation of crystal field theory in detail.

Reformulations in Quantum Field Theory

Experimental and theoretical particle physics is also largely concerned with the calculation of matrix elements, specifically those of the *scattering matrix* (the "*S-matrix*" for short). Elements of the S-matrix are known as *scattering amplitudes*, and they characterize the likelihood that scattering a given set of incoming particles results in a particular set of outgoing particles. As with the matrix element calculations considered above, scattering amplitudes are frequently calculated using perturbation theory. Feynman diagrams provide a traditional method for reformulating the elementary perturbation theory calculation into a much more convenient method for calculating scattering amplitudes. We begin with a Lagrangian representing the dynamics governing the scattering process. Using this, we construct a generating functional for the interacting part of the Lagrangian, which physically represents the probability amplitude for a process that begins and remains in the vacuum state in the presence of interacting fields. The generating functional for the interaction can itself be expressed as a functional derivative of a generating functional for the free-field Lagrangian. This enables us to apply perturbation theory to the free-field generating functional, leading to a Taylor series expansion for the interacting generating functional. Each order in perturbation theory corresponds to a power of the coupling constant(s) between the interacting fields. With this Taylor series expansion in hand, we could naïvely calculate terms in the expansion to our hearts' content. This would serve as an elementary approach to calculating scattering amplitudes.

However, many of the terms in the Taylor expansion do not contribute to the non-trivial, interacting part of the scattering amplitude. The elementary approach is blind to this. Feynman diagrams (and related expressive means) reformulate this Taylor expansion by focusing on whether terms arise from *connected* vs. *disconnected* diagrams. Hence, we can say that a term is connected if it arises from a connected Feynman diagram.[16] For any choice of the Taylor series expansion parameter(s), there are a finite number of connected terms, and only these contribute non-trivially to the scattering amplitude at that order in perturbation theory. Hence, in order to compute the generating functional to a desired order in perturbation theory, we only have to calculate the associated connected terms. More precisely, we can prove that *the generating functional is proportional to the exponential of the sum of all* connected *terms* (Srednicki 2007, p. 65). This result underwrites one of the key epistemic dependence relations standing behind the intellectual significance of Feynman diagrams. We find out that we don't need to know the disconnected terms in order to calculate scattering amplitudes. This modularizes the perturbation theory problem

[16]In position space, we can say that a term is *connected* if there is a path of propagators $\Delta(y-z)$ between every pair of remaining source terms $\int d^4x\, J(x)$ and/or "vertex terms" $g\int d^4x'$. This definition does not mention diagrams.

by eliminating any need to calculate disconnected terms.[17] Unlike many of the previous examples, this epistemic dependence relation comes from topological properties rather than symmetry properties. Section 3.7 discusses this example in detail.

Despite the convenience afforded by Feynman diagrams, they are often not convenient enough for physicists' tastes. In principle, we could calculate any scattering amplitude (for a theory described by a Lagrangian) by using Feynman rules. But in quantum chromodynamics (QCD), the Feynman diagram approach becomes incredibly complicated even for tree-level scattering with only four or five particles. A key reason is that the number of Feynman diagrams can grow very quickly even as the number of particles involved grows slowly. For tree-level QCD gluon scattering, four particle scattering involves four diagrams, five particle scattering has 25 diagrams, six particles require 220 diagrams, and ten particles require more than 1 million diagrams (Elvang and Huang 2015, p. 8). Hence, using Feynman diagrams for processes like these quickly becomes impractical.

This motivates physicists to reformulate the scattering amplitude problem yet again. Modern on-shell recursion methods for calculating amplitudes work extremely well, at least in the high-energy limit where all particles can be treated as massless. These methods reformulate the standard Feynman diagram variables for characterizing scattering amplitudes (helicity, polarization vectors, 4-momenta, etc.) in terms of spinor–helicity variables and ultimately twistors. In conjunction with Cauchy's residue theorem from complex analysis, these variables enable one to develop recursion relations for determining scattering amplitudes with n-many particles in terms of amplitudes with fewer particles. In many theories, it is possible to show that every (tree-level) amplitude can be reduced to simple three-particle scattering amplitudes (Elvang and Huang 2015, p. 62).

These recursion relations are yet another example of modularization: they show that to compute an n-point amplitude, we can modularize the problem into computing lower-point amplitudes. On-shell recursion relations also lead to a diagrammatic representation, known as *blob-diagrams*. Conveniently, a much smaller number of blob-diagrams generally suffices to calculate a scattering amplitude that would require a much larger

[17] In position space, both Feynman diagrams and their corresponding integral terms represent the connectivity properties of the terms in the perturbation series. Hence, there isn't much reason to prefer one over the other. However, in momentum space, Feynman diagrams succeed at representing connectivity properties while the corresponding integral terms fail. This difference in encoding connectivity helps account for the prevalence of momentum-space Feynman diagrams, since it is also more tractable to compute the relevant integrals in momentum space.

number of Feynman diagrams. For instance, some 6-gluon scattering processes require only two blob-diagrams, whereas a Feynman diagram calculation would require 38 diagrams (Elvang and Huang 2015, pp. 34, 59–60). Hence, these on-shell recursion methods provide dramatic simplifications, but they are much more than merely convenient. They completely restructure how we understand scattering amplitudes. Section 6.9 considers how on-shell recursion inspired variable transformations that make a hidden symmetry manifest.

Reformulations in Mathematics

Finally, there are a bounty of examples from pure mathematics that illustrate the intellectual significance of different solution procedures. These involve both alternative methods for computations and alternative proofs of the same theorem. For example, many counting problems can be performed either with elementary combinatorics or by taking advantage of symmetries of the problem—using group actions to apply Burnside's formula. As another computational example, if our goal is to compute 3^{231} modulo 5, Fermat's little theorem teaches us that it is unnecessary to multiply 3 by itself 231 times and find the remainder after dividing by 5.[18] Instead, we can figure out how many multiples of 4 go into 231. These will all be congruent to 1 modulo 5. Hence, we reduce the problem to 3^3 modulo 5. This is easily seen to be 2.

As a third example, we can consider different methods for computing the decimal expansions of irrational numbers (such as pi). Traditional methods work recursively, calculating subsequent decimal places as a function of earlier ones. Yet, mathematicians have in some cases found formulas like the Bailey–Borwein–Plouffe equation, which allows calculating the nth digit of pi (represented in base-16) without knowing the preceding digits.[19] Providing another class of examples, there are many cases where a theorem possesses both algebraic and geometric proofs, and these have a very different intellectual character. Many of these are even relevant for physics. For instance, the Wigner–Eckart theorem can be proved both by focusing on Lie algebras or by focusing on properties of Lie groups. The same holds true for many selection rules and for deriving Clebsch–

[18]Fermat's little theorem states that if p is a prime number and n is a natural number, then n^p is congruent to n modulo p. Provided that p does not divide n, this entails that n^{p-1} is congruent to 1 modulo p. In this case, $3^4 \equiv 1 \pmod{5}$.

[19]Thanks to Gordon Belot for suggesting this example.

Gordan coefficients. One of the advantages of conceptualism is that it provides a uniform treatment of the intellectual significance of both applied and pure mathematical reformulations, all in terms of epistemic dependence relations.

Some of these mathematical cases also place pressure on fundamentalism, now in the context of joints in an imagined platonic space. To take an example from complex analysis, consider Cauchy's integral formula. Given a holomorphic function on an open set in the complex plane that has a smooth simple closed curve γ for boundary, then for any point z_0 in the interior, we have the following relationship:

$$f(z_0) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z - z_0} dz \tag{1.1}$$

This formula provides the following epistemic dependence relation: to know the value of this function at any interior point, it suffices to know the function's values on the boundary γ, provided that we know the function is holomorphic on the region bounded by the curve.[20] There is certainly an intellectually significant difference between calculating $f(z_0)$ this way vs. calculating it directly (and this remains so even if this is not a typical use of Cauchy's formula). Yet, it is difficult to see why one of these ways of formulating the problem should be more joint-carving than the other. Against the conventionalist, it also doesn't seem important whether or not this formula is actually useful. Similar claims hold for the other mathematical examples described above.

Each of these cases illustrates different ways that alternative solution procedures can provide intellectually significant differences. Although the more mathematically sophisticated reformulations frequently provide marked gains in convenience, this seems to be far from the only difference separating them from elementary formulations. Furthermore, at no point have we needed to appeal to a difference in joint-carving to do justice to these reformulations' *prima facie* significance. Remarkably, all of the key aspects of these intellectually significant features can be analyzed in terms of a few generic kinds of epistemic dependence relations: modularization, unification, and uniformity of treatment. There likely other generic kinds of epistemic dependence relations, but it is gratifying to see how much sense we can make of these disparate examples by using so few resources.

As should be clear from these examples, many of their most interesting features come

[20] We also require that the closed contour γ is small enough to exclude other poles, and that there are no branch cuts connected to z_0.

from the content of their particular epistemic dependence relations. For instance, when we modularize the problem of computing scattering amplitudes by figuring out that disconnected terms do not contribute non-trivially, what seems to be particularly significant is the epistemic dependence relation itself: i.e. that *it suffices to compute the connected terms.* Due to the particular nature of each epistemic dependence relation, it is not the case that every instance of modularization is as significant as any other. Some instances of modularization will seem much more significant. I am not optimistic that we can construct a general method for quantifying relative significance, although in some cases it will seem exceedingly clear that a reformulation dramatically restructures our understanding of a theory rather than making a minor intellectual difference. I return to a related question in Section 4.5, where I consider what it means for one formulation to provide a *better understanding* than another. Section 6.4.1 contains one proposal for quantifying the relative significance of intellectually significant changes. We may be able to quantify significance as a function of how many epistemically possible solutions a formulation rules out—at least in comparison with another compatible formulation.

Chapter 2:

Between Instrumentalism and Fundamentalism about Reformulations

2.1 Introduction

Chapter 1 introduced a puzzle about compatible formulations: if no compatible formulation is necessary, what is the value of reformulating a problem-solving procedure or theory? This chapter considers in detail the various solutions to this puzzle introduced in Section 1.3. I begin in Section 2.2 by providing two simple illustrations of compatible formulations. Section 2.3 uses these examples to motivate three desiderata that any satisfying account of reformulations must satisfy. Subsequent sections argue that of the various solutions considered, only my preferred solution—conceptualism—meets these three desiderata.

According to *instrumentalism*, reformulations are valuable insofar as they are a means to the end of traditional scientific goods. These goods include prediction, discovery, control, and descriptive or empirical adequacy.[1] The instrumentalist argues that although reformulations are one method for achieving these aims of science, reformulations themselves are not constitutive of any scientific aims. This kind of instrumentalism has been developed in detail to make sense of the value of scientific explanations (van Fraassen 1980, pp. 93–94; Lombrozo 2011). By itself, instrumentalism leaves certain aspects of reformulation mysterious. It entails that there is no non-instrumental value in having more than one approach to solving the same problem. Yet, it seems that local differences in problem-solving strategy often contribute to different ways of understanding the world. Section 2.4 argues against instrumentalism.

[1]Viewed from within science, these scientific goods have *final* as opposed to instrumental value. As Korsgaard (1983) argues, final value (i.e. non-instrumental value) need not be intrinsic. It need not supervene on the intrinsic properties of the states of affair with final value.

Through additional metaphysical commitments, *fundamentalism* proposes a further scientific aim beyond empirical adequacy or even truth. To fully succeed, science must construct a theory that carves nature at its joints. Whereas Lewis (1983) frames this aim in terms of perfectly natural properties, Sider (2011) has developed a more general notion of the *fundamental structure* described by a theory. According to Sider, successfully describing fundamental structure leads to greater epistemic value. In this way, fundamentalism provides a straightforward account of the value of reformulating. Insofar as reformulating is sometimes constitutive of writing a theory in more joint-carving terms, fundamentalists can interpret some reformulations as non-instrumentally valuable.

For those willing to endorse additional metaphysical commitments, fundamentalism offers a non-instrumentalist account of the value of reformulating. However, it comes at the cost of difficult problems of epistemic access. As I argue in Section 2.5, these epistemic access problems partly spoil the positive story that fundamentalism can tell. Section 2.6 develops conceptualism as a third strategy for assessing the value of reformulations, occupying a middle ground between instrumentalism and fundamentalism. I will argue that reformulations have non-instrumental value simply in virtue of how they restructure problem-solving. Successful reformulations clarify what we need to know to solve problems, improving our understanding of the world. Like instrumentalism, my account does not require substantial ontological commitments. Like fundamentalism, it accommodates the intuition that many reformulations are more than just instrumentally valuable.

2.2 Two Simple Illustrations

As the examples from Chapter 1 intimate, scientific reformulations are often rich and complex. While inherently interesting, such examples require a wealth of background knowledge to assess, as Section 1.5 makes clear. Fortunately, a few simple examples illustrate the general features that arise. To broach the problems facing instrumentalism and fundamentalism, I will focus on two such examples here, the first being even simpler than the second.

Consider the following problem, discussed in the cognitive science literature on problem-solving and expertise (Goldman 1986, p. 132). Two trains, located at stations 50 miles apart, both head toward each other at 25 miles per hour. While they are moving, a bird flies back and forth between them, flying at 100 miles per hour. The problem is to figure out how many miles the bird travels before the trains meet. One relatively *hard approach* to this problem involves calculating the distance the bird flies on each round-trip between the two trains. Stipulating that the bird always takes the shortest distance between the two trains, one can determine the overall distance by summing a geometric series, with a term for each leg of the journey. An *easy approach* to solving this problem involves simply determining how long the bird is in flight. This equals the amount of time it takes for the trains to reach each other, namely, one hour. Hence, the easy approach entails immediately that the bird travels 100 miles as it flies between the trains.[2]

As a second example, consider an application of Gauss's law in electromagnetism. We are handed a ball containing static point charges of total charge Q. Our task is to quantify the strength of the electric field coming out of the ball. In other words, we need to determine the *electric flux* Φ_E, defined as the integral of the electric field $\mathbf{E}$ over the surface.[3] Naïvely, it would seem that to calculate the flux we need to know the electric field vector at each point passing through the surface. And to determine these electric field vectors, it would seem that we need to know the exact distribution of charges within the ball. Incredibly, Gauss's law shows us that we in fact do not need to know anything about either the charge distribution or the electric field to determine the flux. Instead, the electric flux simply equals the total amount of charge contained within our surface divided by a constant ε_0, known as the vacuum permittivity. Hence, knowledge of ε_0 and the total charge Q suffices for knowing the flux.[4]

In both cases, we have two compatible ways of solving the same problem. The procedures do not disagree about the way the world is. They provide the same answer to the problem and ultimately for the same physical reasons, albeit differently organized.

[2]The *mutilated checkerboard problem* provides a similar example: after removing the squares from two opposite corners of a checkerboard, can the remaining squares be tiled with 31 dominoes? For discussion and additional examples, see Bilalić et al. (2019).

[3]More precisely, the electric flux Φ_E through a closed surface S is the surface integral of the component of the electric field normal to the surface, i.e. we integrate the scalar product of the electric field vector $\mathbf{E}$ with the differential of the normal vector to the surface d$\mathbf{a}$: $\Phi_E \equiv \oiint_S \mathbf{E} \cdot \mathrm{d}\mathbf{a}$. Thanks to Gordon Belot for suggesting this example.

[4]Gauss's law also supplies another simple compatible reformulation for systems with appropriate symmetry. In such cases, we can calculate the electric field itself purely algebraically, eliminating the need for integration. In contrast, a non-symmetry-based approach would apply Coulomb's law and a superposition principle for electric fields, integrating for the electric field.

Our question is the following: what value is there to having more than one approach to solving the same problem? More generally, what value is there in reformulating a problem-solving procedure or theory?

Instrumentalism contends that these reformulations are not valuable in themselves, but merely as a means to scientific ends such as better predictions, control, or empirical adequacy. Applied to our two illustrations, instrumentalism entails few differences. In each case, both formulations solve the same problem, so locally we do not have any differences in prediction, control, or (approximate) truth. Each compatible formulation is as good as the other when it comes to obtaining these scientific goods. The only remaining differences between the formulations are practical or pragmatic ones. These include differences in computational simplicity, efficiency, and convenience.

For instance, it is easier and faster to solve the bird–train problem by figuring out how long the bird is in flight than by calculating a geometric series. Likewise, it is easier and faster to apply Gauss's law to determine the electric flux than to painstakingly apply Coulomb's law. The easier methods may in turn decrease the risk of making a calculational mistake, but this is an epistemic difference in-practice, rather than in-principle. Later, I will consider whether global differences in problem-solving fruitfulness allow instrumentalism to draw epistemically significant differences between formulations. Perhaps one formulation generalizes to a wider range of phenomena, leading to increased instrumental value. For reasons considered in Section 2.3, I will argue that differences in fruitfulness still miss important epistemic differences between the approaches.

By contrast, on Lewis's fundamentalist framework, a formulation does better the closer it comes to a canonical language that carves nature at its joints. A concept carves nature *perfectly* at its joints only if it is fundamental, but joint-carving is not an all or nothing affair. Instead, different concepts within the special sciences can be more or less joint-carving (Lewis 1983, p. 347). Sider enriches this picture by arguing that differences in joint-carving generate differences in the epistemic value of formulations. Given two languages for describing the world, if one of them carves nature better at the joints, then it has epistemic value that the other one lacks. Sider illustrates this in the context of the predicates green and grue, claiming that "it's *better* to think and speak in joint-carving terms. We ought not to speak the 'grue' language, nor think the thoughts expressed by its simple sentences" (2011, p. 61).

In the case of the bird and the trains, it is plausible that neither formulation is more joint-carving than the other. The geometric series approach keeps track of the causal details of the bird's trajectory, while the easy approach shows that we do not need this information to solve the problem. Yet, neither approach is obviously more fundamental than the other. In cases like this, a fundamentalist might agree with an instrumentalist that this reformulation has no more than practical value. The Gauss's law case is more interesting. As one of Maxwell's laws of electrodynamics, Gauss's law is plausibly more fundamental than Coulomb's law. Gauss's law is related to conservation principles, which themselves have a close connection with laws of nature and fundamental symmetries (Strocchi 2013, Ch. 7). Additionally, Gauss's law applies to not only static but also moving charges, and it is therefore more general than Coulomb's law. Although this difference in fruitfulness does not directly help with assessing compatible formulations, a fundamentalist might nevertheless view it as evidence that the Gauss's law approach gets closer to fundamental joints in nature.[5]

On the view I defend in Section 2.6, we need not deny that reformulations have instrumental value nor that they could—for all we know—have epistemic value coming from tracking fundamental structure. Instead, we can be sure of one source of their non-instrumental epistemic value: reformulations clarify what we need to know to solve problems. By changing our epistemic situation, reformulations accrue epistemic value independently of any further metaphysical role they might play. In short, a significant reformulation leads to a different way of understanding the world. This is in contrast to trivial or insignificant reformulations, considered in the next section.

2.3 Three Desiderata

I will argue that any satisfying account of reformulations must satisfy three desiderata. First, it must distinguish trivial notational variants from significant reformulations. Whereas some reformulations are merely matters of arbitrary, conventional choices, others appear to be epistemically significant. Second, a successful account must make sense

[5]As Tappenden (2008) notes, defenders of joint-carving (see Section 2.5) may take differences in fruitfulness or fertility as evidence that one formulation is more fundamental than another. Nevertheless, like me in Section 6.6, Tappenden does not endorse a metaphysically robust notion of 'fundamentality.' I remain neutral on whether fruitfulness plays this evidential role, at least when it comes to metaphysically robust notions of 'fundamental.' See also Nolan (1999), who argues that fertility is not a fundamental virtue.

of local differences between reformulations, which arise even when solving the same class of problems. Although reformulations often lead to differences in solving wider classes of problems, appealing to these wider differences alone does not address important local differences. Finally, the criteria that an account employs ought to be epistemically accessible. An account will be less satisfying insofar as it appeals to features of the world that might readily elude us. My goal in this section is to independently motivate these three desiderata. Sections 2.4 and 2.5 will then show how both instrumentalism and fundamentalism fall short of meeting them.

Not all reformulations are epistemically significant. Some amount to nothing more than trivial notational variants. These include simple notational substitutions for typographical preference, the use of a right-handed rather than a left-handed coordinate system, conventions for summation, etc. I take it as a datum of scientific and mathematical practice that these trivial notational variants are epistemically insignificant. At the very least, they are much less epistemically valuable than paradigmatic cases of reformulation, including the two simple cases presented in Section 2.2. A successful account of compatible reformulations must provide principled grounds for distinguishing trivial notational variants from significant reformulations, affording greater epistemic value to the latter. This requirement supplies the *first desideratum*.[6] To satisfy it, an account must avoid both i) over-generating cases of significant reformulations (e.g. by classifying all reformulations as epistemically significant) and ii) under-generating such cases (e.g. by classifying all reformulations as trivial notational variants).

To meet the first desideratum, an account must provide a principled distinction between clear cases of trivial vs. significant reformulations. There may be vague cases that do not fall neatly into either category. Hence, meeting the first desideratum does not require necessary and sufficient conditions for when a reformulation counts as "epistemically significant." It suffices to justify the datum that there is an epistemically significant difference between trivial vs. significant reformulations, with the latter being objectively more epistemically valuable (at least in clear cases).[7] This distinction is objective in the

[6]Readers of Chapter 1 might wonder why I don't state the first desideratum in terms of *intellectual significance*, i.e. objective, non-practical, epistemic differences. I worry that this might unfairly prejudge the issue against instrumentalism. The burden is on me to motivate a distinction between non-practical and practical epistemic features.

[7]Some might be inclined to presuppose that some formulations are objectively epistemically *better than others*. My argument does not depend on this further assumption. I consider the subtle issue of comparative

sense that its truth does not depend on how agents feel or what they believe about it. Regarding the meaning of "epistemically significant" or "epistemic differences," there are many candidates. Different accounts may specify different meanings for these terms. I describe my preferred account in Section 2.6, which focuses on non-practical dimensions of the epistemic.

The second desideratum restricts what we can appeal to when meeting the first desideratum. In clear cases, we can distinguish trivial from significant reformulations at the local level of individual problems or problem-types. This is another apparent datum of scientific practice that any satisfying account must save. Given two compatible reformulations, there is a class of problems that they both solve. Within this shared domain of problems, significant reformulations display an epistemic difference, while trivial reformulations do not. Since these epistemic differences arise locally, we should account for them through local aspects of the formulations. It should not be necessary to consider global differences in fruitfulness or problem-solving scope. Unless shown otherwise, we should assume that these global differences arise from differences at the local level of solving individual problems. The second desideratum embodies these demands: a satisfying account of reformulations must provide local criteria for distinguishing trivial vs. non-trivial reformulations. Section 2.4 will show how this desideratum poses a serious problem for instrumentalism.

Besides the need to locally distinguish trivial from significant reformulations, a *third desideratum* presents itself: the criteria of significance should be epistemically accessible. Since the epistemic difference between trivial vs. non-trivial reformulations is manifest, the criteria we use to explicate this difference should be manifest as well. Our account of reformulation should not be hostage to the lucky success of risky inferences. An account with epistemically inaccessible criteria may have the resources to address the first two desiderata, but it would be difficult to determine when the criteria actually apply. Accounts of reformulation that rely on risky inferences will face problems of underdetermination, leading to skeptical scenarios. The more difficult it is to determine whether the criteria are satisfied, the more skeptical scenarios there will be. In science, these worries about underdetermination are well-founded: there are numerous historical examples of

assessments between compatible formulations in Sections 4.5-4.6.

scientists making needlessly risky inferences that were shown to be unfounded.[8] This is not an idle philosopher's skepticism. There are principled, practice-based reasons for seeking to avoid risky inferences whenever possible.

An additional dialectical consideration speaks in favor of the third desideratum as well. Appraising compatible formulations is a challenge facing philosophers of many different temperaments, from hard-nosed scientific anti-realists to those willing to posit Aristotelian essences. Ideally, an account of reformulation should have a widely-acceptable minimal core. This core should be as minimal in its ontological commitments or posits as possible. Nothing precludes those with additional metaphysical commitments from embellishing this account further, but it is harder to deconstruct a more metaphysically committal account into a version acceptable for the a-metaphysical. Section 2.5 will show how this third desideratum severely limits the appeal of fundamentalism, at least as the core of an account of reformulations.

2.4 Problems facing Instrumentalism

To distinguish trivial from non-trivial reformulations, instrumentalism proposes the following criterion: a reformulation is significant if and only if it leads to an instrumentally valuable difference. What counts as instrumentally valuable might vary from case to case. Some reformulations might make a difference for prediction, others for control, and still others for empirical adequacy or truth. Since these kinds of instrumental differences matter epistemically[9], instrumentalism can in principle satisfy the first desideratum. Instrumental differences can provide objective, epistemic grounds for distinguishing trivial notational variants from significant reformulations.

However, this strategy for meeting the first desideratum runs afoul of the second. When we focus on how two compatible formulations solve *the same* class of problems, we see that there are often no significant instrumental differences to be found. Since the two formulations locally solve the same problems, they locally provide the same predictions, are of equal approximate truth, etc. Hence, it is difficult to see how there could

[8]Examples include Newton's inference from absolute acceleration to the existence of absolute velocity and position, 18th-century inferences to the existence of caloric as the carrier of heat, and 19th century inferences to the existence of an ether for the propagation of light as an electromagnetic wave.

[9]If prediction and control are instead non-epistemic, practical features of science, then this would make it harder for instrumentalism to satisfy the first desideratum.

be local differences in instrumental epistemic value beyond differences in convenience. At most it seems, one formulation might be more convenient for solving problems than another, perhaps involving simpler calculations or more familiar notation. As I argue in Section 2.4.1, differences in convenience are insufficient to distinguish trivial from non-trivial reformulations.

For instrumentalism to meet the first desideratum, it seemingly must appeal to global differences, such as differences in fruitfulness. When we broaden our scope to consider how reformulations differentially generalize in different contexts, sometimes certain formulations succeed where others fail. For instance, the easy solution to the bird–train problem applies even to a bird executing exquisite loop-de-loops between the trains. In contrast, the geometric series solution requires that the bird fly in straight lines (otherwise, we would require further empirical information about the bird's trajectory). Similarly, the Gauss's law approach applies to moving charges, while the Coulomb's law approach requires that the charges are static. In each case, one formulation is more fruitful than the other, applying to a strictly wider range of problems.

No doubt, differences in fruitfulness are instrumentally valuable. They constitute differences in the predictions we can make and the phenomena we can save. However, they are not differences that arise at the local level of shared problem-solving. To appeal to them alone would be to give up on the goal of accounting for local differences that are *prima facie* significant. We should expect that these global differences are symptoms of underlying local differences in problem-solving. Thus, it would be more satisfying if we could accommodate global differences in terms of local, epistemically significant differences between formulations. We should give up on the second desideratum only if other promising accounts fail to meet it as well. For this reason, instrumentalism is not enough to account for the significance of reformulations. Instrumental differences are part of a larger story, but they are not the whole story.

2.4.1 Convenience-instrumentalism

At this point, an instrumentalist might want to revisit local differences in convenience as a criterion for meeting the first and second desiderata. Why can't these successfully distinguish trivial from non-trivial reformulations? There are at least two serious problems facing what we might call *convenience-instrumentalism.*

First, paradigmatic cases of trivial notational variants often display important differences in convenience. For instance, we find it extremely difficult to read mirror images of words.[10] Similarly, scientists sometimes develop such a strong psychological preference for certain notational conventions that they find it difficult to solve problems using alternatives. For instance, the 1959 English translation of Wigner's (1931) text on the applications of group theory to atomic spectra converted Wigner's original left-handed coordinate system into a right-handed convention to facilitate comprehension. Similarly, physicists working in particle physics phenomenology tend to use a different space-time metric convention than those working in general relativity or string theory. The former tend to use a mostly minus $(1, -1, -1, -1)$ metric while the latter use a mostly plus $(-1, 1, 1, 1)$ metric. Although just a choice of convention, "some physicists approach this issue with almost religious conviction" (Burgess and Moore 2006, p. 518). There are many compelling practical reasons to prefer one convention over the other, based on the kinds of problems that most commonly arise in either domain. Convenience-instrumentalism would entail that these two metric conventions are significant reformulations. Such verdicts would vastly over-generate the class of significant reformulations—by under-generating the class of trivial notational variants—thereby running afoul of the first desideratum. Even if these practical differences between trivial notational variants are sometimes important, they still appear to be *different in kind* from the epistemic differences between significant reformulations.

The second problem afflicting convenience-instrumentalism stems from the subjective nature of convenience. What counts as computationally simpler or more convenient is often—if not always—a matter of taste and pedagogical training. For instance, comparative assessments of convenience can depend on the tools at one's disposal. Whether one finds it simpler to use (i) a mathematically sophisticated formulation that reduces problems to arithmetic vs. (ii) an elementary formulation that requires calculating integrals might depend on whether one has access to numerical integration software. It also depends on one's level of mathematical background, which influences preferences for abstract vs. elementary mathematics. As the early history of applied group representation theory shows, the lack of familiarity with abstract mathematics can lead to a preference

[10] For an illustration, see Wittgenstein (2009 [1949], p. 209), remark 151 of *Philosophy of Psychology*. Framing effects from the presentation of statistics in terms of decimals or ratios provides a further example. See Kahneman (2011).

for methods that involve elementary but tedious calculations.[11] Insofar as assessments of convenience are subjective, the first desideratum precludes convenience as a criterion of significance. For recall that the first desideratum requires an objective distinction between clear cases of trivial vs. non-trivial reformulations. To capture this distinction, the corresponding criteria must be objective.

A third concern arises as well: differences in convenience are plausibly practical differences rather than epistemic ones. They amount to differences in how easy it is to solve a problem, rather than differences in reasons for belief. However, since I have so far left unspecified what counts as an epistemic difference, it is difficult to develop this objection in a non-question-begging way. It is open to the convenience-instrumentalist to argue that practical differences in increasing calculational speed or decreasing the risk of error are also epistemic differences. After all, they might note, the use of computers in science have led to both a rapid expansion in scientific knowledge and a decreased risk of calculational error. Regardless, I argue in Section 2.6.3 that we can meaningfully distinguish practical from non-practical epistemic differences.

For now, I take the preceding two problems as sufficient grounds for rejecting convenience-instrumentalism. It does not provide a satisfactory account of the objective epistemic value of significant reformulations. A convenience-instrumentalist would have to provide both an objective account of "convenience," and an error theory for ordinary scientific assessments of trivial vs. non-trivial reformulations. Unless these challenges are met, we should reject convenience-instrumentalism for failing to save central aspects of scientific practice. It is a philosophical position of last resort, to be adopted only if alternative accounts face equally challenging problems.

2.5 Problems facing Fundamentalism

According to many scientific realists, science aims at the truth. Fundamentalism proposes a further aim for science: an ideal scientific theory must describe the world in a fundamental language. Two descriptions of the world can both be true, while one of them is more fundamental. Lewis contends that physics aims at providing an inventory of natural properties (1983, p. 357). According to Lewis, "the businesss [sic] of physics is not just to

[11] For historical accounts, see Bonolis (2004) and Scholz (2006).

discover laws and causal explanations. In putting forward as comprehensive theories that recognize only a limited range of natural properties, physics proposes inventories of the natural properties instantiated in our world" (1983, p. 364). Likewise, Sider argues that describing the world in joint-carving terms leads to greater epistemic value than merely having a true theory:[12]

> The goal of inquiry is not merely to believe truly (or to know). Achieving the goal of inquiry requires that one's belief state reflect the world, which in addition to lack of error requires one to think of the world *in its terms*, to carve the world at its joints. Wielders of non-joint-carving concepts are worse inquirers. (2011, p. 61)

Although neither Lewis nor Sider are explicitly concerned with compatible formulations, their commitments to fundamental structure suggest a *fundamentalist criterion* for assessing reformulations: a reformulation is epistemically valuable whenever it leads to a more joint-carving formulation. Using this criterion, fundamentalism straightforwardly meets the first two desiderata from Section 2.3. It proposes an objective epistemic difference between trivial notational variants and significant reformulations. Whereas trivial notational variants are equally joint-carving, significant reformulations exhibit a difference in fundamentality: namely, one formulation is more joint-carving than the other. Furthermore, the fundamentalist criterion of significance is local: these differences in fundamentality arise at the level of individual problem-solving. Fundamentalism thereby satisfies the second desideratum as well. Although evidence for differences in joint-carving might come from global considerations of fruitfulness, the differences themselves arise locally (if they arise at all).

The main problems facing fundamentalism arise from its substantial ontological commitments. Many empiricists and scientific anti-realists (and even some realists) disavow commitments to perfectly natural properties and fundamental structure. Relying on these commitments precludes fundamentalism from providing a minimal account of the value

[12] Dasgupta (2018) has challenged Sider's contention that more fundamental descriptions necessarily have greater epistemic value. He argues that fundamentalists must *explain* where this epistemic value comes from, but that no explanation is forthcoming. However, I worry that Dasgupta's argument is self-undermining. Fundamentalists fail to meet Dasgupta's demand for an explanation only if we presuppose that a realist conception of explanation is desirable. This amounts to presupposing a value claim little different than what the fundamentalist is accused of presupposing. Hence, Dasgupta's own argument is subject to either circularity or an infinite regress. The fundamentalist can simply demand that Dasgupta explain why the fundamenalist owes an explanation of the epistemic value of approximating fundamental structure. By his own lights, Dasgupta will not be able to meet this demand, so his own demand for explanation is self-undermining.

of reformulations. In response, a fundamentalist might be inclined to say: so much the worse for the metaphysically-averse. But there are independently compelling reasons to be wary of appeals to fundamental structure. One reason comes from fundamentalism itself: Occam's razor. If we can provide a positive account of reformulations with fewer metaphysical commitments, then this account will be simpler. Fundamentalists would then, by their own lights, have reasons to take such an account seriously.[13] This is one reason in favor of the conceptualist account I provide in Section 2.6.

More substantial metaphysical commitments typically involve posits that are less epistemically accessible. It is difficult to know if and when theory formulations track perfectly natural properties. Beyond appeals to intuition, fundamentalists must rely on theoretical virtues as evidence for greater fundamentality. Whether and when these differences in virtues—such as simplicity or fruitfulness—themselves track fundamentality is an issue that is itself ultimately decided by appeals to philosophical intuition. Some scientific realists and fundamentalists may be sufficiently optimistic about these aspects of philosophical methodology. For them, these epistemic access problems may not be substantially more troubling than our access to the physical unobservables posited by ordinary scientific theories. Nevertheless, an account of reformulations would be epistemically more secure if it did not rely on these controversial methodological commitments.[14] Ideally, we should seek an ontologically minimal account of reformulations that even empiricists can adopt. More metaphysically-committed philosophers then remain at liberty to invoke additional ontological commitments when assessing reformulations.

Epistemic access problems also lead to problems of underdetermination. Imagine a world where neither the Gauss's law nor Coulomb's law formulation is more fundamental than the other (see Section 2.2). This world is empirically indistinguishable from the one fundamentalists might take ourselves to be in, where the Gauss's law formulation is putatively more fundamental. In either world, our physical theories make exactly the same predictions about both observables and unobservables. Yet, the two worlds disagree about whether the Gauss's law formulation is more fundamental, and hence about whether the two formulations are trivial notational variants. In the former world, the fundamentalist

[13] Thanks to Dave Baker for stressing this point.

[14] Cohen and Callender (2009, p. 13) argue that perfectly natural properties face additional problems of epistemic access beyond the usual skeptical challenges to knowledge of physical unobservables. See Woodward (2016, p. 1056) for additional criticisms against simply invoking natural properties.

criterion of significance classifies the two formulations as trivial notational variants. In the other, this criterion says that the formulations are significantly different. But since both worlds are empirically indistinguishable, it is difficult to know which one we are in.[15]

As a result, fundamentalism makes the significance of compatible formulations hostage to empirically inaccessible facts about fundamental structure. Even worse, these inaccessible facts do not have any bearing on how the formulations appear to us. Metaphysical facts about differences in joint-carving do not change how we solve problems or understand the world using our theories.

To avoid underdetermination problems, we should strive for an account of compatible formulations that does not depend on relatively inaccessible facts about the world, e.g. its fundamental structure. Even if we are mistaken or woefully ignorant about the world's fundamental structure, we should be able to satisfactorily interpret significant epistemic differences between reformulations. In contrast, fundamentalist differences between reformulations are not epistemically accessible, making it ambiguous whether and when there are significant differences between formulations. Fundamentalism meets the first two desiderata *in principle*. But in virtue of failing the third desideratum, fundamentalism makes it difficult to know when significant differences arise. To avoid this problem, we need an account that makes the epistemic differences between formulations manifest. The next section develops one such account.

2.6 Conceptualism

Having seen the problems facing instrumentalism and fundamentalism, I now develop an account of reformulations that avoids these problems. Because it focuses primarily on how different concepts restructure problem-solving, I call my account *conceptualism*. Conceptualism accounts for the significance of reformulations in terms of how they organize or structure problem-solving. Section 2.6.1 provides examples of these organizational differences. They constitute differences in the epistemic structure of a problem-solving procedure. Section 2.6.2 shows how conceptualism easily satisfies all three desiderata, providing a positive, local, and ontologically minimal account of reformulations. Finally,

[15]van Fraassen (1975) develops a similar underdetermination problem for mathematical Platonism through his parable of the lands of Oz vs. Id. Cohen and Callender (2009) provide this kind of argument against the epistemic accessibility of perfectly natural properties.

Section 2.6.3 rebuts an objection that my account fails to distinguish trivial from significant reformulations.

2.6.1 Epistemic structure

Consider the toy example of the bird flying between two trains from Section 2.2. The hard procedure requires knowing the distance the bird travels on each leg between the trains (or alternatively, the time spent on each leg). The easy procedure shows that we do not need to know the bird's detailed trajectory: it suffices to know the speed at which the bird flies and the amount of time spent flying. Similar considerations apply to calculating the electric flux emanating from a charged body using Gauss's law. Gauss's law shows that we do not need to know the distribution of charges within the charged object or the electric field at each point on the surface. Instead, it suffices to know the total amount of charge the object contains.

These simple cases of reformulation parallel more complicated cases. Using symmetry arguments, we can often solve problems without needing to know many details about a system's dynamics. In quantum chemistry, reformulating eigenvalue problems in terms of molecular symmetries shows that we don't need to know much about particular energy levels to solve problems about overall energy structure and allowed transitions. In contrast, elementary methods that eschew appeals to symmetries require detailed information about the molecular energy levels. Similarly, in atomic and nuclear physics, the Wigner–Eckart theorem shows that to calculate the expectation values of physical quantities, we don't need detailed information about each possible state. Since these states are related by symmetries of the system, it suffices to calculate one process in detail. This provides the symmetry-invariant content for each symmetry-related family of states. The Lagrangian reformulation of classical mechanics illustrates a similar moral. It tells us that to calculate the equations of motion for a classical system, we do not need to know the constraint forces acting on the system. In contrast, the Newtonian formulation requires knowledge of these constraint forces. I will argue that these reformulations are significant because they provide epistemically different ways of solving the same class of problems.

In each of these cases, reformulating a theory or problem-solving procedure results in an alternative epistemic structure. The epistemic structure of a solution procedure comprises the inferential steps we take in moving from a given set of inputs to an output.

For each of these steps, there are things we need to know—or at least that suffice to know—to move from input to output. In short, this is the information we use to carry out an inference rule. For convenience, I will refer to these input–output relations as *epistemic dependence relations* (EDRs). EDRs specify what we need to know or what suffices to know to carry out a step in a problem-solving plan. Generically, EDRs state whether we need to know fact B in order to know fact C, or whether fact D is sufficient for deriving fact E. Relations such as these serve as basic units of a theory's epistemic structure.

I will often focus on EDRs that relate the initial inputs to the final output, i.e. the solution. These EDRs characterize what we *need to know* or what *suffices to know* in order to solve a problem. In this context, we are considering what we need to know *tout court*.[16] Yet, an exclusive focus on these kinds of EDRs would neglect the internal structure of a problem-solving plan. This internal structure manifests what we need to know (or what suffices to know) to carry out a particular inference rule. Hence, many questions of what we need to know to solve a problem are implicitly relativized to an inference rule that we intend to use.[17] In contrast, when we consider what we need to know *tout court* to solve a problem, our question ranges over any admissible inference rule.

Differences between Arabic and Roman numerals provide a simple illustration of epistemic dependence relations. Imagine a lecture hall with 21 rows of 16 seats each. Our task is to determine how many people it can seat. Arabic numerals allow us to multiply 16 by 21 using a standard algorithm from grade school. This algorithm takes advantage of Arabic numeral's positional notation to modularize the problem into a series of single-digit multiplication and addition sub-problems, such as calculating six times two. Thanks to this algorithm, to solve any multiplication problem, it suffices to know (i) 100 single-digit multiplication facts (i.e. the times table up to 9) and (ii) how to add Arabic numerals. Now, imagine reformulating this multiplication problem using Roman numerals, i.e. calculating XVI times XXI. Since Roman numerals are a sign-value system rather than a positional

[16] As Chapter 5 discusses in detail, the minimal success criteria for solving a problem are different for constructive empiricists vs. scientific realists. For the former, approximate empirical adequacy is sufficient. For the latter, approximate (non-deflationary) truth is needed. Hence, when it comes to unobservables, claims about what an empiricist needs to know are elliptical for "what they need to know according to the theory." Empiricists do not think that we need to know any claims about unobservables in order to solve scientific problems, since they claim that we don't even need to believe claims about unobservables. Here, I intend to develop conceptualism as a neutral schema that either constructive empiricists or scientific realists could adopt as a starting point.

[17] I thank Laura Ruetsche for prompting me to characterize EDRs more generally so as to capture this internal structure.

one, our familiar algorithm does not work.[18] We must rely instead on different epistemic dependence relations. Rather than an addition table, we instead use seven simplification rules such as replacing "IIIII" by "V". We also use a multiplication table of 49 separate multiplication facts (such as L times L equals MMD), which must be augmented for factors above one million.[19] When it comes to figuring out that 16 times 21 equals 336, these two formulations display different epistemic structures, characterized by their different EDRs. They amount to different plans for problem-solving.

2.6.2 Satisfying the three desiderata

Conceptualism proposes a straightforward criterion for assessing the epistemic significance of reformulations: two formulations are significantly different when they provide different epistemic dependence relations. As I now argue, this criterion satisfies all three desiderata from Section 2.3.

The first desideratum demands a principled distinction between trivial notational variants and significant reformulations. Unlike significant reformulations, trivial notational variants fail the above criterion. When we trivially reformulate a theory or problem-solving procedure, we preserve epistemic structure. Trivial notational variants provide the same epistemic dependence relations for solving problems.[20] Symbol substitution provides the simplest case: substituting every instance of a symbol α with a previously unused, arbitrary symbol β does not alter the formulation's epistemic dependence relations. Likewise, even though many scientists prefer to work in a right-handed coordinate system, working in a left-handed coordinate system preserves the same EDRs. In relativistic theories, the arbitrary choice between a mostly positive or a mostly minus metric convention does not lead to differences in epistemic dependence relations. Hence, these two conventional choices are trivial notational variants. This is the case despite the fact that many physicists have a personal—and sometimes subfield-wide—preference for one convention over the other. After discussing the other two desiderata, I will return to whether

[18] In an additive system, the string represents the sum of its individual numerals. For simplicity, I do not consider subtractive notation such as "IV" for four, representing this instead as "IIII." Everything I say below could be adapted for this case. See Detlefsen et al. (1976).

[19] See Schlimm and Neth (2008, p. 2100) for details of this algorithm, which relies on the distributive law. In Roman numerals, figuring out which of two numbers is greater also involves different EDRs in these two formalisms; see Colyvan (2012, pp. 133–134).

[20] Framed in terms of syntactic symmetries, the epistemic structures/plans provided by trivial notational variants are invariant under the reformulation. Section 2.6.3 discusses this interpretation in more detail.

trivial notational variants in fact exhibit different EDRs.

The conceptualist criterion also satisfies the second desideratum, which demands that we *locally* distinguish trivial from significant reformulations. Differences in epistemic structure arise at the local level of solving individual problems. We can assess whether two compatible formulations are significantly different by considering their shared class of problems. We need not appeal to differences in their fruitfulness or scope. Differences in fruitfulness are no doubt also epistemically significant, but conceptualism shows how they arise from local differences in epistemic dependence relations. It is in virtue of restructuring our solution procedures that some formulations become more fruitful than others for certain classes of problems. Differences in fruitfulness are not a reason for significance; they are a symptom.

Finally, conceptualism satisfies the third desideratum, which calls for epistemic accessibility. Unlike fundamentalism, conceptualism does not require inductively-risky inferences to fundamental ontology. Even anti-realists about physical unobservables can recognize differences in epistemic dependence relations. These differences between reformulations are typically manifest. We learn about EDRs simply by analyzing how formulations support problem-solving. For instance, when we discover a new way to solve a problem, we learn that we didn't need to know certain facts that we previously relied on. This is not to say that we have complete knowledge of all relevant EDRs. Typically, reformulations show that a particular set of facts *suffices* for problem solving, or that an alternative set of facts is *not necessary*. Conceptualism merely points out that successful reformulations implicitly provide proofs of this form. Proving that a set of facts is *insufficient* or that a set of facts is *necessary* would require additional proofs, which are not typically implicit in reformulations themselves.

In virtue of our relatively easy access to epistemic structure, conceptualism avoids the underdetermination problems that afflict fundamentalism. Even in a world where we are radically wrong about which formulation is more fundamental, we will be right about many differences in epistemic dependence relations. These differences in EDRs are not empirically underdetermined. In contrast, fundamentalism relies on differences in fruitfulness or other super-empirical virtues as *evidence* for deeper ontological differences. This involves making an epistemically risky *inference* to the existence of underlying differences in fundamental structure. Conceptualism provides a method for appraising reformulations that avoids these risky inferences.

Both scientific realists and fundamentalists may hanker for a deeper explanation of epistemic dependence relations. They may seek to ground these EDRs in explanatory differences or differences in fundamental structure. For instance, perhaps differences in EDRs correspond to differences in what information is explanatorily relevant to the problem-solution. According to this explanationist proposal, information that *we do not need* to solve a problem is always explanatorily irrelevant. In the case of the bird and the trains, we do not need to know the detailed trajectory that the bird takes. Some may therefore be inclined to say that this detailed trajectory is explanatorily irrelevant. One difficulty with this inference is that it takes us from considering the epistemic structure of a formulation to considering more contentious explanatory relations in the world. Philosophers who support causal-mechanical accounts of explanation may have a different intuition. From a causal-mechanical standpoint, the bird's detailed trajectory explains the distance it travels. It remains explanatorily relevant, despite the fact that we do not need to know it in order to solve certain problems. Conceptualism shows that we can positively assess reformulations without resolving these kinds of explanatory disputes. Section 2.7 considers explanationism in more detail.

Nothing prevents philosophers with a more optimistic view of theoretical virtues from making further inferences about physical or metaphysical facts that ground EDRs. They are welcome to do so if so inclined. Nevertheless, these additional commitments preclude fundamentalism and some forms of explanationism from providing an ontologically minimal account of reformulations, based on epistemically accessible resources. If instrumentalism could meet the first two desiderata, it would already provide a minimal account. But as it stands, instrumentalism is inadequate. At the other extreme, fundamentalism commits us to more than necessary. Conceptualism, I have argued, is just right. Of course, there may be other philosophical accounts of reformulation that satisfy these three desiderata. I would be surprised if any such account proves to be deeply incompatible with conceptualism.

2.6.3 Sameness of epistemic structure

For conceptualism to satisfy the first desideratum, trivial notational variants must have the same epistemic structure. This requires that they invoke the same epistemic dependence relations. But how do we know when two formulations have the same EDRs? What is the criterion for *sameness* of EDRs and, more generally, *sameness* of epistemic structure? Here, I will defend a simple account of sameness of epistemic structure. Much more could be said, since the issue parallels synonymy of meaning (of which there is a vast literature). However, for my purposes it suffices to provide a plausible criterion for distinguishing different epistemic structures. My account relies on an equivalence between plans, specifically plans for problem-solving. This will show that it is possible to draw an intelligible and non-question-begging distinction.

Characterizing *sameness of epistemic structure* initially seems difficult for the following reason: even obvious cases of trivial notational variants require knowing slightly different relations (in some sense of "different"), simply because they involve different notation. For instance, to solve a problem using a left-handed coordinate system, you need to understand the relevant convention. This left-handed convention is *a fortiori* different than that of a right-handed convention. To take a linguistic analogy, knowing what "dogs bark" means requires knowing some English, while knowing what the synonymous expression "die Hunde bellen" means requires knowing some German. It seems clear that these kinds of conventional differences do not yield differences in meaning (in the relevant sense). Typically, notation is a vehicle for communicating content, not the content itself. Yet, if one simply stipulates that differences in convention or linguistic knowledge are not epistemically significant, there is a risk of begging the question. What counts as an epistemically significant difference is exactly what is in question here.

Fortunately, it cannot be the case that *any change* in notation suffices for an epistemically significant change. Otherwise, there would be no such thing as trivial notational variants. It would then be impossible to understand a sense in which "dogs bark" means the same as "die Hunde bellen." Clearly, there is sense to be made of this claim of synonymy.[21] Insofar as we take there to be trivial notational variants, we take there to be some sense in which the epistemic structure of a formulation is not fully determined by the notation we use to express it. When we say that "a dog cannot dissemble," we assert something independently of the particular language we voice it in. We assert the same thing as what we could voice in German as "ein Hund kann nicht heucheln." Of course,

[21]Even if one has sympathies for Quine's skepticism about meaning and synonymy, it remains true that ordinary assertions of synonymy often *seem* intelligible.

saying the latter requires knowing some German, but knowledge of German is not essential for understanding the content of this sentence. Although understanding any sentence requires understanding a notation, that does not make the notation part of the content of the sentence.

Following Gibbard's (2012) account of meaning, we can understand the synonymy of "dogs bark" and "die Hunde bellen" as follows. Both sentences voice the same thought, which we can denote either as DOGS BARK or DIE HUNDE BELLEN. To believe the sentences are synonymous (in a given situation) is simply to plan to use "dogs bark" if I am an English speaker in those situations that I would plan to use "die Hunde bellen" if I were a German speaker, and vice versa. Hence, synonymy of meaning is simply a matter of synonymy of plans. Although the English and German speaker know different languages, once we extract away these linguistic differences, they know the same thing, namely the thought that DOGS BARK.

Applying Gibbard's account of synonymy to epistemic structure yields a clear criterion for trivial notational variance. Imagine that one could solve a given problem with either a left-handed or a right-handed coordinate convention. In the left-handed case, I appeal to an EDR expressed in the left-handed convention, denoted 'EDR_{left}.' In the right-handed case, I appeal to an EDR expressed in the right-handed convention, denoted 'EDR_{right}.' EDR_{left} and EDR_{right} voice the same EDR provided that when working in a left-handed convention, I plan to use EDR_{left} in the same situations as I would plan to use EDR_{right} if I were working in a right-handed convention. Hence, although I technically need to know something different to work with EDR_{left} rather than EDR_{right} (and vice versa), this difference is an artifact of my notation, rather than a genuine epistemic difference in my problem-solving plan.

To clarify further, let's consider a simple example worked out in detail. For many problems, it does not make an epistemic difference whether we express percentages as decimals or fractions. Instead, relative to many problem-solving procedures, these two notations are trivial notational variants: they voice exactly the same EDRs (albeit expressed in their respective notations). We see this in the following kind of simple problem: you are given a quantity and asked whether it is less than 5%. In the decimal formulation, you are presented with '0.04.' You know the following epistemic dependence relation, $\text{EDR}_{decimal}$: to convert a decimal to a percent, it suffices to multiply by 100 and affix a percent symbol.

In practice, this operation may involve moving the decimal point two slots to the right, arriving at 4%. In the fraction formulation, you are presented with '4/100.' You know the following epistemic dependence relation, $\mathrm{EDR}_{fraction}$: to convert a fraction to a percent, it suffices to multiply by 100 and affix a percent symbol. In practice, this operation may involve canceling out the 100 in the denominator by the 100 we are multiplying, again arriving at 4%. Despite their different notation and written operations, $\mathrm{EDR}_{decimal}$ and $\mathrm{EDR}_{fraction}$ voice the same EDR.

In each formulation, what I plan to do at each step of the problem-solving procedure matches what I plan to do in the other formulation, up to the notational differences in how I voice these inferences and carry them out on the page. Thus, these two expressive means are trivial notational variants, relative to this problem-solving procedure. In typical problem-solving contexts, similar morals apply to our choice of space-time metric convention (mostly minus vs. mostly plus) or our choice to use Einstein summation convention (vs. explicitly writing 'Σ_i' for each index i we sum over).

Unlike trivial notational variants, significant reformulations provide different plans for solving problems. Ultimately, this is borne out as a difference in the epistemic dependence relations that they exploit or make available. For example, in the bird–train problem, someone using the hard formulation needs to determine the distance the bird travels on its first segment, second segment, etc (or, alternatively, the time spent on each segment). They then need to know how to sum the distance on these segments, relying on an EDR for summing an infinite geometric series. An agent following the easy formulation does not need to determine this information, nor rely on this EDR. This is a genuine difference in epistemic structure, i.e. in problem-solving plan.

Making a Property Manifest

Although it is natural to speak about particular notations as being trivial notational variants or not, this is actually elliptical for whether two problem-solving procedures are trivially different or not. Two notations can be trivial notational variants with respect to one class of problems but non-trivial with respect to another. We see this already in the case of decimals and fractions. If our problem is to determine how many people out of 1000 have a certain medical condition, then presenting an incidence rate in fraction form can be epistemically different from presenting the same quantity in decimal form.

If you are told that the incidence rate is 6/1000, you can immediately infer that 6 out of every 1000 people have this condition. In contrast, if you are told that the incidence rate is 0.006, you need to know how to convert this decimal to a fraction out of 1000. The notation "6/1000" *makes manifest* the number of cases out of 1000. The notation wears the answer on its sleeves; no further EDR is required.

Making a property manifest is a more subtle kind of epistemic difference that arises in many examples, such as polar vs. Cartesian coordinates, hidden vs. manifest symmetries, planar representations of planar graphs, etc. Despite the existence of a translation procedure from one notation to the other, there can be differences in when we are licensed to make certain inferences—based on the notation we are using. For instance, if you were told that the incidence rate is 6000/1000000, you again need to know how to convert this to a fraction out of 1000, e.g. by dividing by 1000. Chapter 6 discusses how some formulations succeed at making a property manifest, while others—even those related by a translation—do not.

Hence, the existence of a translation procedure between notations does not entail that they are always trivial notational variants. What matters is whether the notations ever support problem-solving procedures with different epistemic structures, i.e. epistemically different plans. If they do, then they are—in that context—non-trivial notational variants. This can happen even if there is a uniform way to translate from one notation to another. Trying to exploit a particular EDR in one notation might require translating to the other. Whereas if we had started out in this latter notation, we would not need to translate.[22]

This subtle kind of epistemic difference parallels how certain linguistic meanings can be "lost in translation." For instance, imagine that i) when solving a problem in English you either need to look up a word in an English dictionary or translate a word to German, but ii) if you were solving the same problem in German you would not need to do either. In a case like this, the German formulation of this problem is epistemically different, even if each German word can be translated to an English word. For instance, a German can reliably guess that "die Speisekarte" voices the concept MENU, since "die Speise" means 'dish' or 'food' and "die Karte" means 'card' or 'chart.' In contrast, the meaning of the English word "menu" is not manifest from knowledge of the English words "dish" and "card." Section 6.3 elaborates this example.

[22] See Sections 6.2.1-6.5 for a few detailed illustrations.

An analogous phenomenon arises in the context of mathematical isomorphisms. Two mathematical objects can be isomorphic in one sense, but not another. For instance, two objects can be isomorphic as vector spaces, but not as Lie algebras, or isomorphic as groups but not as rings. To be isomorphic as vector spaces means that these objects support the same plans for solving a particular class of problems. We could translate between the objects for this class of problems and use either one for problem-solving. Yet, the existence of a translation procedure in one context does not entail that this procedure holds for every possible problem-solving context (e.g. problems in the context of Lie algebras, where the vector isomorphism does not entail synonymy of problem-solving plans in this new context).

Practical vs. Non-Practical Epistemic Differences

Focusing on sameness of epistemic structure also allows us to distinguish non-practical epistemic differences from practical differences that can arise from notational choices. For instance, I might have a strong preference to work in a right-handed coordinate convention, to work with fractions rather than decimals, or to read ordinary rather than mirror-image text. I might be considerably faster with one notation than the other. Indeed, I might even be more reliable working with one notation than the other, making fewer mistakes. Some of these practical differences, especially in terms of speed or reliability, might strike some philosophers as being genuinely epistemic.[23] It is of course fine to call them 'epistemic' if one prefers. What matters is that they are importantly different from the kinds of epistemic differences that significant reformulations provide. Such differences in problem-solving plans exist independently of anyone's preferences, comfort-level, speed, or risk of error. They exist even for ideal computers (although not necessarily for logically-omniscient agents). These non-practical epistemic differences are what Section 1.2 calls "intellectually significant." Practical differences in reliability or speed between trivial notational variants are not epistemic in the same sense, even if one wishes to call them epistemic in some other sense.[24]

To take another example, consider how our beliefs are often subject to practically

[23]See, for instance, Goldman (1986, 122ff.), who discusses speed as an epistemic standard for assessing intelligence.

[24]Chapter 5 further distinguishes between (i) practical differences in speed or convenience and (ii) non-practical, intellectually significant differences. See especially Sections 5.6-5.7.

important framing effects. Cognitive psychology has found evidence that people are more likely to believe that the risk of disease is high when presented as a proportion, i.e. number of people out of a reference population size (such as 5 out of 1000). People are more likely to believe risk is low when the same information is presented as a percentage (e.g. 0.5%). The former makes salient that actual people have this condition, and that you could be one of those people. The latter does not, since percentages do not make the number of people salient.[25] For instance, saying that 1 out of every 500 Americans have died from COVID sounds much worse than the equivalent 0.2%. Despite being practically important for rhetorical purposes, this is not an intellectually significant difference. What one ought to believe about the risk does not depend on how the proportion is expressed.

A similar point holds for differences in problem-solving speed between formulations: what one ultimately ought to believe does not depend on how quickly the problem can be solved, or how resource-intensive its solution is. Of course, if one computer can get you a solution in five minutes, and the other takes two weeks, that is practically important for belief-formation. However, if the two computers implement the same solution procedure (e.g. program)—just on different hardware—then there is no intellectually significant difference between them.

To summarize, two problem-solving procedures have the same epistemic structure if agents following either procedure ought to believe the same information, step-by-step. This means that an agent following one procedure ought to plan to make the same inferences as an agent following the other. Even though these procedures might be written in different object languages, such differences are genuinely notational: they do not impact the thoughts or content voiced.[26] Sameness of epistemic structure amounts to synonymy of problem-solving plans. Trivial notational variants therefore display a symmetry of epistemic structure: the epistemic structure of the problem-solving procedure is invariant under the change in notation. This symmetry does not relate the notations themselves; rather, it relates the problem-solving procedures written down in these notations. Hence,

[25] For discussion, see Kahneman (2011, Ch. 30). Even worse, humans are biased by the size of the numerator (neglecting the denominator), even when comparisons are both made in fraction form: people rated a disease that killed 1286 out of 10000 people as riskier than one which killed 24.14 out of 100 people, even though the former is half as deadly as the latter (Yamagishi 1997).

[26] Of course, we assume that the agents begin in evidentially-equivalent starting points. Although they might have knowledge of different languages or notation, these differences are intellectually insignificant. Gibbard's (2012) account of synonymy provides a detailed framework for explicating this intuitive condition while avoiding vicious circularity.

two notations might be trivial variants in one problem-solving context but not in another. This is again analogous to isomorphism of mathematical structures. Whereas philosophical discussions of theoretical equivalence typically focus on equivalence of theories or models, here we have focused on equivalence and non-equivalence of problem-solving procedures. It is through these problem-solving procedures that we actually come to understand the world (or other subject matters such as mathematics).

2.7 Problems with Explanationism

I have argued that conceptualism meets the three desiderata laid out in Section 2.3. It provides a middle ground between instrumentalism and fundamentalism about reformulations. Of course, there might be other intermediate positions that meet these three desiderata as well. *Prima facie*, one approach that seems attractive would involve tracking putative differences in explanation. Perhaps two compatible formulations are epistemically different provided that they exhibit explanatory differences. I will call this schematic proposal *explanationism*. It satisfies the first desideratum by holding that trivial notational variants are explanatorily on a par, whereas significant reformulations manifest explanatory differences. Provided that these explanatory differences are local and epistemically accessible, explanationism will meet the second and third desiderata as well. In this section, I argue that conceptualism has important advantages over explanationism. In particular, conceptualism characterizes the epistemic differences between reformulations without taking a stand on the contentious topic of explanation.

Whether or not two compatible formulations have an explanatory difference depends on the nature of explanation. Different accounts of explanation give diametrically opposed verdicts on the simple examples that we have considered. Hempel's (1965) deductive–nomological account treats both the easy and hard approaches to the bird–train problem as equally explanatory: both appeal to the same law-like statement (distance as a function of rate and time), the same initial conditions, and provide equally rigorous derivations of the explanandum. Hence, a Hempelian explanationist would have to view these as trivial notational variants. On a causal–mechanical account of explanation, the hard approach to the bird–train problem might be viewed as more explanatory, since it explicitly tracks additional causal details. Likewise for the Coulomb's law approach to calculating electric flux, since this approach explicitly calculates the electric field from each individual charge. A unificationist account of explanation suggests the opposite verdict: by eliminating reference to these additional causal details, the simple approach to the bird–train problem and the Gauss's law approach both apply to a wider range of phenomena.[27] Despite combining aspects of causal and unificationist approaches, Strevens' (2008) kairetic account of explanation would agree with the unificationist verdict here. The kairetic account claims that whenever we can abstract information from a causal model—while still saving the phenomena—this information is explanatorily irrelevant.

These disagreements illustrate the important role that philosophical assumptions play in assessing what information qualifies as explanatorily relevant. In contrast, we can recognize that certain reformulations do not require information that another formulation requires, without needing to make further philosophical assumptions about explanation. Hence, we can more securely discern that reformulations display epistemic differences than explanatory differences. Moving from a recognition of these epistemic differences to claims about explanatory relevance requires further philosophical principles.

For instance, when we find out that knowledge of the distance traveled on each leg of the bird's journey is unnecessary for solving the bird–train problem, we might be tempted to infer that this information is explanatorily irrelevant. Doing so requires endorsing a philosophical principle like the following: contextually-unnecessary but causally efficacious information is explanatorily irrelevant. Proponents of causal–mechanical pictures of explanation may have different philosophical intuitions about whether this contextually-unnecessary information is explanatorily irrelevant. They might instead argue that tracking this information provides a deeper explanation, even if this deeper explanation is unnecessary for many purposes. My point here is a simple one: settling this sort of philosophical dispute is downstream from characterizing central epistemic differences between compatible formulations. We can account for many of the epistemic and methodological advantages of reformulations without settling these further questions about explanation or explanatory relevance.

Most accounts of explanation agree on at least one aspect of explanation: explanations

[27]For this traditional dialectic between causal-mechanical vs. unificationist account of explanation, see Salmon (1998).

provide answers to why-questions.[28] Explanatory information describes the reasons why an event occurred or a fact is true. This aspect of explanation provides a second argument for viewing explanatory differences as logically downstream from the epistemic differences that concern conceptualism. Logically, why-questions form a proper subset of a larger category of scientific and mathematical questions. Not all problems take the form of requests for explanatory information or reasons why. Hence, not all problem-solving procedures provide explanations, even if they succeed at providing solutions. Questions about whether a solution procedure is explanatory typically go beyond whether it provides the correct solution. We see this, for instance, in the case of mathematics: a rigorous proof of a mathematical theorem may not count as explanatory. For instance, Lange (2009b) argues that proofs by mathematical induction often fail to be explanatory.

In privileging conceptualism over explanationism, I do not deny that philosophical questions about explanation and explanatory relevance are important. My point is merely that various versions of explanationism could agree with my conceptualist analysis of reformulations, while disagreeing about the nature of explanation. Conceptualism is therefore much better suited to provide the minimal core of an account of reformulations. Having adopted this minimal core, one can then defend further philosophical principles about explanation and explanatory relevance. In this way, my complaint against explanationism is similar to my complaint against fundamentalism: to assess important epistemic differences between reformulations, explanationism has to presuppose more than necessary. Chapters 3 and 4 develop additional arguments against various versions of explanationism, which I take to be the most compelling alternatives to conceptualism.

2.8 Conclusion

Ultimately, the value of reformulations comes from how they facilitate the aims of science. An instrumentalist account of reformulations fits well with empiricist-friendly aims such as prediction, control, and empirical adequacy. Reformulations can then be seen as instrumentally valuable for bringing about these scientific goods. Yet, instrumentalism lacks adequate resources for distinguishing trivial from significant reformulations. Fundamentalism endorses a more metaphysically substantial aim for science, namely to arrive at a

[28]This includes both irrealist frameworks such as van Fraassen's (1980) pragmatic account and realist approaches such as Skow's (2016) causal–grounding account of reasons why.

fundamental language for describing reality. It therefore has the resources to interpret some reformulations as more than merely instrumentally valuable. According to fundamentalism, significant reformulations constitute more fundamental descriptions of reality. However, this positive account comes at the cost of greater ontological commitments, leading to worries about underdetermination.

Conceptualism provides a middle ground between instrumentalism and fundamentalism. It preserves the positive features of these accounts, while avoiding their drawbacks. Understood in terms of the aims of science, conceptualism augments empiricist-friendly aims with a further one: scientists should aim at clarifying the epistemic structure of their theories. They should seek to determine what they need to know to solve scientific problems. Just as empirical adequacy is a scientific good, so is the clarification of epistemic structure.[29] By figuring out what we need to know to solve problems, we enhance our understanding of the world, independently of any further downstream benefits such as greater fruitfulness, better explanations, or more fundamental descriptions. Significant reformulations lead to different epistemic dependence relations. They thereby constitute a clarification of epistemic structure. Provided this clarification is intellectually significant, so is the reformulation that contributes to it. In this way, conceptualism can interpret reformulations as having non-instrumental, epistemic value. Advantageously, conceptualism's positive account does not require any special-purpose ontological commitments. It thus provides a minimal core for a positive appraisal of reformulations.

[29]Chapter 5 argues that the clarification of epistemic structure is a *constitutive aim* of science. Meeting this aim is part of the minimal success criteria for ideal scientific theorizing.

Chapter 3:

Understanding and Equivalent Reformulations

3.1 Introduction

Accounts of theoretical equivalence have neglected an important epistemological question about reformulations: how does reformulating a theory change our understanding of the world? *Prima facie*, improving our understanding is one of the chief intellectual benefits of reformulations. Nevertheless, accounts of theoretical equivalence have focused almost entirely on developing formal and interpretational criteria for when two formulations count as equivalent (Weatherall 2019). Although no doubt an important question, focusing on it alone misses many other philosophically rich aspects of reformulation.

The burgeoning literature on scientific understanding would seem to be a natural home for characterizing how reformulations improve understanding. However, existing accounts of scientific understanding do not provide a clear answer. These accounts tend to focus on *competing* rather than *compatible* explanations, investigating how the best explanation provides understanding. This strategy neglects how equivalent formulations of the *same explanation* can provide different understandings. To address these gaps, I will show how theoretically equivalent formulations can change our understanding of the world.[1]

Harkening back to Hempel, Kitcher, and Salmon, the *received view* of understanding holds that understanding why a phenomenon occurs amounts to grasping a correct explanation of that phenomenon (Strevens 2013; Khalifa 2017, p. 16). Many recent accounts of understanding have decried this picture as overly simplistic, arguing that genuine understanding goes well beyond grasping an explanation (Grimm 2010; Hills 2016; Newman

[1]Sections 3.3-3.6 of this chapter were published in Hunt (2021b).

2013; de Regt 2017). Nevertheless, in trying to augment the received view, these critics still maintain a close connection between explanation and understanding. Khalifa (2012, 2013, 2015) has exploited this connection to systematically undermine their more expansive accounts. Defending what I'll call *explanationism*, Khalifa (2017) has argued that we can reduce understanding-why to the epistemology of scientific explanation. Explanationism thereby poses a serious challenge to accounts of understanding that seek to go beyond the received view.

Here, I argue that we can refute explanationism by considering theoretically equivalent formulations. By definition, theoretically equivalent formulations agree completely on the way the world is, thereby describing the exact same state of affairs. Moreover, philosophers often adopt an *ontic conception* of explanation, wherein explanations themselves correspond to states of affairs, e.g. the reasons why an event occurs.[2] By agreeing on the way the world is, equivalent formulations *ipso facto* provide the same explanations. Nonetheless, they can differ radically in the understandings that they provide. Thus, concerning many phenomena, theoretically equivalent formulations do not differ qua explanation, even as they differ qua understanding. These differences in understanding—without concomitant explanatory differences—make a separate account of understanding necessary.

I begin in Section 3.2 by clarifying my target: explanationism. Next, Section 3.3 summarizes Khalifa's explanationist challenge for existing accounts of scientific understanding, showing how they reduce to accounts of explanation. I focus in particular on how Khalifa problematizes both skills-based accounts of understanding and a different strategy developed by Lipton (2009) that foreshadows my own. Section 3.4 demonstrates that theoretically equivalent formulations provide a large class of cases that meet Khalifa's challenge. In these cases, we have differences in understanding-why without differences in explanation. In Section 3.5, I defend *conceptualism* as a positive account of these differences in understanding. Conceptualism characterizes how such differences can arise from the presentation and organization of explanatory information. It meets Khalifa's challenge while accommodating the significance of reformulations. Section 3.6 consid-

[2]For the ontic conception, see Salmon (1998 [1984], p. 325), Strevens (2008, p. 6), Craver (2014), and Skow (2016). Non-ontic conceptions of explanation also often have an ontic component. My positive proposal shows one way of assessing how reformulations change our understanding, without needing to assume or defend an alternative pragmatic or epistemic conception of explanation.

ers and rebuts an objection to my use of theoretically equivalent formulations. Finally, Section 3.7 provides an extended illustration of my proposal in the context of Feynman diagrams.

3.2 Explanationism

Philosophers who work on scientific understanding sometimes allege that traditional theories of explanation neglected scientific understanding, or at least did not say enough about it (de Regt 2017, p. 16). Nonetheless, these philosophers typically agree with traditional accounts of explanation on a central schema that connects understanding with explanation. According to this "received view of understanding," understanding why a phenomenon occurs consists in grasping an explanation of that phenomenon.[3] By tightly connecting understanding-why with explanation, the received view transforms even traditional accounts of scientific explanation into a minimal account of scientific understanding.

The received view suggests two sources for differences in understanding why a phenomenon occurs. First, on the agentive side, these differences can spring from variation in how agents grasp explanations. Most recent accounts of scientific understanding have focused their attention here, arguing that understanding involves special skills or abilities for grasping explanations.[4] These agentive aspects of understanding-why are ideally intersubjective but often idiosyncratic.[5] Second, on the non-agentive side, differences in understanding can arise from grasping different explanatory information, such as different states of affairs or other ontic features of reality. Differences in ontic explanatory features straightforwardly provide objective and non-pragmatic (i.e. 'non-practical') differences in understanding. Insofar as traditional accounts of explanation have been in-

[3]For statements of this position see Strevens (2013), Khalifa (2017, pp. 16–8), de Regt (2017, p. 23), and Potochnik (2017, pp. 123–4).

[4]See, for instance, de Regt and Dieks (2005), de Regt, Leonelli, et al. (2009), Grimm (2010), and Hills (2016). Along with scientists' cognitive abilities, Potochnik's account of understanding relies on scientists' research interests, background information, space-time location, and psychological characteristics (2017, p. 100).

[5]de Regt (2017, p. 44) argues that the pragmatic nature of skills does not entail that the resulting understanding is problematically subjective. Potochnik makes a similar claim regarding her account of understanding, where "features of scientists themselves, including their interests and intentions, influence what generates understanding" (2015b, p. 74). However, these agentive features seem less objective than both explanatory differences in understanding and the epistemic differences that Chapter 2 considers.

terested in scientific understanding, it has been in this second sense.

This traditional focus leads to *explanationism*, which claims that all objective and non-practical differences in understanding arise from differences in the ontological content represented or picked out by explanations.[6] Phrased as a biconditional, explanationism contends that an intellectually significant difference occurs if and only if there is a corresponding ontic explanatory difference. Such ontic differences comprise differences in the worldly features responsible for the phenomenon of interest, such as laws of nature, causes, mechanisms, grounds, and difference-makers. Explanations that appeal to different ontic features straightforwardly lead to intellectual differences; this establishes one direction of the explanationist biconditional. Many accounts of explanation also treat such ontic differences as necessary for an intellectual difference, establishing the second direction.

For instance, according to Hempel, "all scientific explanation [...] seeks to provide a systematic understanding of empirical phenomena by showing that they fit into a nomic nexus" (1965, p. 488). Similarly, Trout argues that the only kind of understanding that we should focus on is an objective kind coming from explanations, namely "the state produced, and only produced, by grasping a true explanation" (2007, pp. 584–5). Strevens defends this same claim (2008, p. 3), arguing further that "science understands a phenomenon just in case it can provide a standalone explanation of the phenomenon," namely "an explanation that is complete, that is not missing any of its parts" (2008, p. 117). Woodward also frequently makes remarks that are congenial to explanationism, such as his claim that "once we have been given information about the complete patterns of counterfactual dependence [...] it appears that nothing has been left out that is relevant to understanding why matters transpired as they did" (2003, p. 86). On this traditional conception, non-ontic differences—such as differences in the mode of presentation of an explanation—are seen as being practical or pragmatic.

As a thesis about the relationship between understanding and explanation, explanationism is not itself an account of explanation. As such, there are a great variety of explanationists, distinguished by their preferred accounts of scientific explanation. This includes defenders of both *ontic* and *epistemic* conceptions of explanation. As characterized

[6] What I am calling "explanationism" might more precisely be called "ontic explanationism." It is distinct from weaker positions seeking to reduce intellectual differences to both ontic and non-ontic features of explanation.

by Salmon, the "ontic conception" views explanations as objective and non-pragmatic features of the world that exist independently of explanatory arguments or discoveries (1989, p. 133). In this ontic sense, an explanation is "a relation among features of the world," namely the features "that cause, produce, or are otherwise responsible for the phenomena we seek to explain" (Craver 2014, pp. 30, 36). In contrast, the "epistemic conception" of explanation focuses on the representation of these ontic features. The epistemic conception privileges explanation-texts or explanatory arguments as being explanations proper. Nevertheless, as Craver (2014) has argued, these explanation-texts must still refer to ontic explanatory information in order to distinguish explanations from non-explanations. Thus, at least for my purposes here, the debate between epistemic and ontic accounts is mainly terminological.[7] For instance, the three approaches to crystal field theory discussed in Chapter 4 provide the same explanation in an ontic sense, but of course they each provide a different explanation-text (and hence a different epistemic explanation). What matters is that they agree on the ontic explanatory information, and it is immaterial if we characterize this information within an ontic vs. an epistemic conception of explanation. Hence, I intend to argue against any account of explanation that takes grasping ontic explanatory information as necessary and sufficient for objective, non-pragmatic differences in understanding.

Pragmatic or agentive accounts of understanding also aim to challenge explanationism, but they are dialectically less effective for this purpose. Following Elgin (2004), Potochnik (2017, p. 95) rejects the traditional factivity assumption that understanding requires truth, requiring instead that the relevant claims be "true enough." According to Potochnik, whether a scientific claim is true enough to provide understanding depends partly on pragmatic considerations, including "the purpose of the research to which it contributes" (2017, p. 96). Ultimately, Potochnik extends these pragmatic considerations to explanation, arguing that the audience "helps determine the nature of the explanatory facts, that is, the ontic explanation" (2017, p. 128).[8] However, it is unlikely that explanationists would willingly grant the assumptions of a framework where ontic explanation depends on features of agents. In general, explanationists are simply less interested in

[7]Even defenders of the epistemic conception, such as Bokulich, agree that "ontic constraints still play a central role" in explanation (2018, p. 794).

[8]Morrison (2000, pp. 28–9) makes a similar claim that the explanatory power or acceptability of an explanation depends partly on the scientific community or even individual scientists.

more subjective, pragmatic conceptions of explanation or understanding.[9]

Pragmatic accounts of understanding face an additional problem, developed in the next section. Khalifa (2012) has noted that accounts of explanation already implicitly involve the use of skills. Agents obviously require some cognitive abilities to construct and grasp explanations. For instance, Woodward's manipulationist account of causal explanation implicitly references the relevant skills for constructing and analyzing what-if-things-had-been-different questions. Thus, explanationism already seems compatible with skills-based accounts of understanding. In contrast, I intend to rebut explanationism on its own terms by privileging its preferred sense of understanding. I will argue that explanationism is incomplete even with regards to these objective and non-pragmatic differences in understanding. Whereas pragmatic accounts of understanding criticize explanationism for reasons it might not find compelling, conceptualism points out a shortcoming that even explanationists should regard as important.

3.3 The Challenge from Explanationism

By closely connecting understanding with explanation, traditional accounts of explanation lead to a deflationary stance toward understanding. According to Khalifa, "on the old view, if understanding was not merely psychological afterglow, it was nevertheless redundant, being replaceable by explanatory concepts without loss" (2012, p. 17). Explanationism encapsulates this deflationary position:

> *Explanationism*: all philosophically significant aspects of understanding-why are encompassed by an appropriately detailed account of the epistemology of scientific explanation.[10]

Importantly, even non-deflationary accounts of scientific understanding must adopt some account of scientific explanation. Then, given whatever account of explanation is adopted, explanationism demands an argument that understanding-why does not reduce to claims

[9]Hempel (1965, pp. 425–32) discusses pragmatic features of explanation at length, noting that although they are important, they can be separated from his non-pragmatic account of explanation and understanding. He remarks that "to propound those [non-pragmatic] models is therefore neither to deny the pragmatic 'dimension' of explanation nor to belittle its importance" (1965, p. 426).

[10]In earlier work, Khalifa refers to this position as the *explanatory model of understanding* (2012, p. 17). Khalifa (2017, p. 85) uses "explanationism" in a narrower sense aimed at showing how objectual understanding can be reduced to explanatory understanding, ultimately defending what he calls "quasi-explanationism." For convenience, I simplify this more cumbersome terminology.

about (this kind of) explanation. For this reason, explanationism is dialectically most effective when married with explanatory pluralism (Khalifa 2017, p. 8). Then, no matter which account(s) of explanation is ultimately correct, explanationism challenges non-deflationary accounts of understanding on their own terms.

Khalifa defends explanationism by developing the *explanation-knowledge-science* (EKS) model: an agent improves their understanding why *p* provided that they either (a) gain a more complete grasp of *p*'s *explanatory nexus* or (b) their grasp of this explanatory nexus comes closer to *scientific knowledge* (2017, p. 14). The *explanatory nexus* is the "totality of explanatory information about *p*," comprising all correct explanations of *p* and the relations between them (2017, p. 6). In Section 3.4, I will argue that knowledge of this nexus does not exhaust differences in understanding-why. Khalifa argues that *scientific knowledge* arises from a three-step process of *scientific explanatory evaluation* (SEEing), involving i) considering plausible potential explanations, ii) comparing these potential explanations, and iii) deciding how to rank these potential explanations with respect to approximate truth (2017, pp. 12–13). Khalifa uses SEEing to deflate many anti-explanationist accounts of understanding.

The primary anti-explanationist strategy argues that understanding-why involves special skills or abilities. Provided these skills go beyond what's required for knowledge-why, explanationism would be refuted.[11] Versions of this *skills-based* strategy include skills for grasping counterfactual information (Grimm 2010, 2014), "cognitive control" over providing and manipulating explanations (Hills 2016), and inferential skills used in making certain kinds of models (Newman 2013). de Regt has provided one of the most sustained defenses of the skills-based strategy, arguing that understanding involves the ability to make qualitative predictions using an intelligible theory that explains the phenomenon (de Regt and Dieks 2005; de Regt 2009, 2017).

Khalifa's criticism of Grimm succinctly illustrates explanationism in action. Khalifa argues that Grimm's (2010) account of understanding makes no advance over Woodward's (2003) account of explanation. According to Grimm, understanding is an ability to predict how changing one variable changes another variable, *ceteris paribus* (2010, pp. 340–41). Yet, as Khalifa notes—and Grimm acknowledges (2010, p. 341, 2014, p. 339)—this

[11]Some have pursued other strategies, arguing that objectual understanding either does not reduce to understanding-why or else that objectual understanding does not require explanatory understanding. Khalifa responds at length to these approaches (2017, p. 80).

kind of understanding is closely related to Woodward's analysis of "what-if-things-had-been-different questions." Hence, this kind of counterfactual reasoning ability is part of scientific explanatory evaluation (SEEing). We already deploy counterfactual reasoning in considering and comparing alternative explanations, and explaining already involves the ability to answer these what-if questions (Khalifa 2017, pp. 71, 74). Khalifa's response is easily generalized: if all that a theory of understanding adds is referencing a cognitive ability to use an explanation, then a theory of explanation can make the same move without modification.[12]

A distinct anti-explanationist strategy seeks cases of scientific understanding in the absence of explanations. Such cases would seemingly show that accounts of explanation miss something about understanding. Undertaking precisely this strategy, Lipton (2009) considers a number of cases where we acquire the cognitive benefits of explanations without actually providing explanations. These cognitive benefits include knowledge of causes, necessity, possibility, and unification (2009, p. 44). Against the received view, Lipton identifies understanding with "the cognitive benefits that an explanation provides" rather than with "having an explanation" (2009, p. 43). This maintains a close connection between understanding and explanation.

Khalifa (2013) exploits this connection to argue that Lipton's strategy makes no fundamental advance over the explanation literature. Systematically examining each of Lipton's examples, Khalifa shows that whenever there is understanding through a non-explanation, there is an explanation that provides *that* understanding *and more*. This leads to "*explanatory idealism*" about understanding, which holds that "other modes of understanding ought to be assessed by how well they replicate the understanding provided by knowledge of a good and correct explanation" (2013, p. 162). Thus, a suitably detailed account of scientific explanation would provide the same insights about understanding that Lipton defends. In this way, explanation functions as the "*ideal* of understanding" (Khalifa 2013, p. 162).

The remainder of this chapter defends a strategy that avoids Khalifa's objections against existing accounts of scientific understanding. My strategy succeeds where

[12]Khalifa (2012) applies this strategy to criticize de Regt and Dieks (2005) and de Regt (2009) in detail. Against Hills, Khalifa argues that her necessary conditions for understanding are either irrelevant for enhancing understanding or else are captured by the EKS model (2017, pp. 70–72). He responds to Newman in his (2015).

others fail for two reasons. First, I do not rely on positing any special abilities unique to understanding, so Khalifa's challenge from SEEing does not apply. Second, the examples I consider provide understanding through the same explanatory information, so explanatory idealism does not apply either.

3.4 Intellectual Differences without Explanatory Differences

To refute explanationism, it suffices to identify differences in understanding-why between two presentations of the *same* explanation, since these appeal—*ipso facto*—to the same explanatory information. In such cases, understanding-why still arises from an explanation, but non-explanatory differences account for the corresponding differences in understanding. The features we ascribe to "understanding-why" and to "explanation" then truly come apart. For convenience, I refer to objective, non-practical differences in understanding as *intellectual differences*. This section discusses cases of intellectual differences without concomitant explanatory differences.

To forestall a piecemeal explanationist response, my argument requires a sufficiently large class of examples stemming from scientific practice. As we will see, the recent literature on theoretical equivalence provides a rich set of cases, spanning many parts of physics. Nevertheless, some might worry that these mathematical reformulations are too isolated or special to be indicative of scientific understanding in general. Hence, it is worthwhile to also consider a more common aspect of scientific practice: diagrammatic reformulations. I will consider both cases in turn, illustrating each with a paradigmatic example.[13] Importantly, my argument does not apply to cases of different but *complementary* explanations, such as Salmon's example of causal-mechanical vs. unificationist explanations of a balloon moving forward upon takeoff in an airplane (Salmon 1998, p. 73; de Regt 2017, p. 77). Such complementary explanations appeal to different explanatory information and are hence genuinely different explanations. Khalifa's EKS model of understanding accommodates such cases since they reference different parts of the explanatory nexus (2017, p. 25).

By definition, *theoretically equivalent formulations* express the same scientific theory, agreeing exactly on the way the world is (or could be). Intuitively, two formulations

[13]Reformulations of symmetry arguments provide another class of examples; see Chapter 4.

are theoretically equivalent if and only if they are mutually inter-translatable and empirically equivalent. Mutual inter-translatability requires that anything expressed in one formulation can be expressed in the other without loss of physically significant information. Empirical equivalence requires that the formulations agree on all physically possible measurable consequences.

Recent defenses of categorical equivalence have shown it to be a fruitful criterion for theoretical equivalence. It successfully formalizes a number of philosophically and scientifically plausible cases of theoretically equivalent formulations.[14] Five prominent examples include Lagrangian and Hamiltonian formulations of classical mechanics (Barrett 2019), standard and geometrized formulations of Newtonian gravity theories (Weatherall 2016), Lorentzian manifold and Einstein algebra formulations of general relativity (Rosenstock, Barrett, et al. 2015), Faraday tensor and 4-vector potential formulations of classical electromagnetism (Weatherall 2016), and principal bundle and holonomy formulations of Yang–Mills gauge theories (Rosenstock and Weatherall 2016). Here, then, is a varied class of cases that collectively pose a serious problem for explanationism.

In each of these cases, I contend, we have intellectual differences without corresponding explanatory differences. Each formulation provides a different understanding than its equivalent counterpart for at least the following simple reason: understanding one does not entail understanding the other (and indeed, showing that they are equivalent requires non-trivial insights). For instance, understanding a phenomenon via Lagrangian mechanics does not entail an understanding of that same phenomenon using Hamiltonian mechanics. Thus, Lagrangian understanding-why differs from Hamiltonian understanding-why, even though both involve grasping the same explanation. The lack of explanatory differences follows from categorical equivalence, which entails that we can inter-translate models of one formulation into models of the other without losing any information.[15] In other words, equivalent formulations possess "the same capacities to represent physical situations" (Rosenstock, Barrett, et al. 2015, p. 315). On the common ontic conception of explanation assumed here, explanatory information itself is a subset of this physical information, so equivalent formulations *a fortiori* represent the same explanatory information.

[14] For an introduction see Halvorson (2016, p. 601) and for details Weatherall (2016, 2019).

[15] For defenses of this claim, see Weatherall (2016, pp. 1083, 1087) and Rosenstock, Barrett, et al. (2015, p. 314). My argument in Chapter 4 has the advantage of not presupposing that categorical equivalence is a good standard for theoretical equivalence.

Thus, whenever one formulation provides an explanation, any equivalent formulation provides the same explanation, preserving everything of ontic explanatory significance—but not necessarily of intellectual significance.

Lagrangian and Hamiltonian mechanics provide a simple but detailed illustration of the foregoing points.[16] These equivalent formulations display two main sources of intellectual differences. First, they differ in how they encode the system's dynamics. The Lagrangian formalism uses a *Lagrangian function* $L(q_i, \dot{q}_i, t)$, encoding the dynamics as a function of time t, generalized coordinates q_i, and generalized velocities $\dot{q}_i$. In the Hamiltonian formalism, we perform a variable change from generalized velocities to generalized momenta p_i, yielding the Hamiltonian $H(q_i, p_i, t)$. Despite encoding the same physical information, the Lagrangian and Hamiltonian organize this information differently, as illustrated below. Second, the two formulations represent the dynamical laws of evolution (the equations of motion) in dramatically different ways. Whereas the Lagrangian formulation represents these as a set of n-many *2nd-order* differential equations (the Euler–Lagrange equations), the Hamiltonian formulation represents these same equations of motion as a set of $2n$-many *1st-order* differential equations (Hamilton's equations).[17] By reorganizing the equations of motion in this way, the Hamiltonian formulation treats the generalized coordinates q_i and the generalized momenta p_i more symmetrically. This leads to further intellectual differences in cases like the following.

A typical explanandum in mechanics concerns the evolution of a classical system such as a pendulum or spinning top. In systems with symmetry, one generalized coordinate, e.g. q_n, is typically *ignorable*—meaning that it does not occur in the Lagrangian or Hamiltonian. The equations of motion then entail that the corresponding conjugate momentum, p_n, is a conserved quantity, i.e. a constant α. It is here that a dramatic intellectual difference occurs between the formulations. Despite p_n being constant, the corresponding generalized velocity $\dot{q}_n$ need not be. Hence, $\dot{q}_n$ still appears in the Lagrangian as a non-trivial variable. A Lagrangian understanding of the system's evolution thereby still requires considering n-many degrees of freedom, despite having an ignorable coordinate. In contrast, the Hamiltonian formalism enables a genuine reduction in the number of de-

[16]Technically—within a subclass of models known as the hyper-regular domain—Barrett (2019) shows that the Lagrangian tangent bundle and Hamiltonian cotangent bundle formulations are equivalent. For ease of exposition, I present their more elementary coordinate-based formalisms. For details see Goldstein et al. (2002).

[17]In both cases, we require $2n$ initial values to solve these equations.

grees of freedom that need to be considered, resulting in a different understanding. By changing variables from generalized velocities to generalized momenta, the Hamiltonian depends on the latter but not the former. Hence, we can replace p_n in the Hamiltonian with a constant α, and—since the ignorable coordinate q_n is absent—this eliminates an entire degree of freedom from consideration.[18] As Butterfield remarks, this example "illustrates one of mechanics' grand themes: exploiting a symmetry so as to reduce the number of variables needed to treat a problem" (2006, p. 43). Although not an explanatory difference, this variable reduction demonstrates a difference in how the same explanatory content is organized. This organizational difference results in a different understanding of the system's evolution. Indeed, these kinds of organizational differences ultimately lead to differences in understanding Noether's first theorem—a foundational result connecting continuous symmetries and conserved quantities (Butterfield 2006).

Thanks to their rigorous mutual inter-translatability, categorically equivalent formulations provide the most precise illustration of my argument. However, at a less rigorous level, theoretically equivalent formulations can arise whenever we reformulate a theory while keeping its physical content the same. This motivates including at least some instances of diagrammatic reasoning within the class of theoretically equivalent formulations. Although neglected by the literature on theoretical equivalence, diagrammatic reformulations satisfy the same intuitive criteria: mutual inter-translatability and empirical equivalence. They thereby provide another large class of examples where we can have differences in understanding-why without concomitant explanatory differences. Examples of diagrammatic reformulations include Feynman diagrams in particle and condensed matter physics, graphical approaches to the quantum theory of angular momentum (Brink and Satchler 1968), Penrose–Carter diagrams in space-time theories, graph-theoretic approaches to chemistry (Balaban 1985; Trinajstić 1992), and diagrams for mechanistic reasoning in biology (Abrahamsen and Bechtel 2015).

To illustrate how diagrammatic reasoning can provide intellectual differences, consider Feynman diagrams in particle physics. Here, the explanandum is typically a scattering amplitude for a particular interaction, explained by calculating terms in a perturbation expansion. Without using Feynman diagrams, we can calculate each term up to a desired order in perturbation theory. This provides one way of understanding the scatter-

[18]Technically, we replace one of Hamilton's equations with a trivial integral for calculating $\dot{q}_n$.

ing amplitude. Alternatively, we can reorganize this same explanatory information using Feynman diagrams, allowing us to express *connectivity properties* of terms in the perturbation expansion. To calculate the scattering amplitude, it suffices to know the connected terms; the disconnected terms contribute only to the non-interacting part of the scattering amplitude, i.e. the identity.[19] Focusing on connectivity thereby makes it unnecessary to consider a vast number of terms in the perturbation expansion—terms that a brute force calculation would show contribute only to the identity. In this way, Feynman diagrams lead to a different understanding of scattering amplitudes but without introducing any additional explanatory information.[20] Section 3.7 discusses Feynman diagrams in detail.

3.5 A Conceptualist Account of Understanding

I have argued that a variety of mathematical and diagrammatic reformulations provide intellectual differences without associated explanatory differences. If not from explanatory differences, whence do these intellectual differences arise? Conceptualism provides a satisfying answer, showing how intellectual differences can result from differences in the organization of explanatory information. These organizational differences generate differences in *what we need to know* to present explanations, leading to differences in understanding-why. I will consider an objection that conceptualism merely describes how reformulations modify explanatory concepts, with no effect on understanding-why. To rebut this objection, I will argue that non-trivial changes in explanatory concepts necessarily lead to differences in understanding-why.

Conceptualism posits a sufficient condition for differences in understanding-why: reformulating an explanation generates an intellectual difference whenever it changes what we *need to know* or *what suffices to know* to present that explanation. For instance, in shifting from Lagrangian mechanics to Hamiltonian mechanics, we learn that we don't need to know how to represent the system and its dynamics using the Lagrangian and the Euler–Lagrange equations. Knowledge of the Hamiltonian and Hamilton's equations suffices. *Mutatis mutandis*, the same can be said for shifting from Hamiltonian mechanics

[19] In position space, a term is connected if there is a path of propagators connecting every pair of source factors and/or vertex factors in the term. For details, see Section 3.7.

[20] de Regt (2017, p. 251) also considers Feynman diagrams to defend his account of understanding, focusing on how they aid visualization.

to Lagrangian mechanics, leading again to a difference in understanding. Similarly, reformulating scattering amplitude explanations using Feynman diagrams teaches us that we don't need to know the disconnected terms in the perturbation expansion: knowledge of the connected terms suffices. For convenience, I refer to these differences in what-we-need-to-know or what-suffices-to-know as *epistemic dependence relations* (EDRs). Conceptualism claims that when equivalent formulations provide different epistemic dependence relations, they manifest intellectual differences.

To rebuff explanationism, these intellectual differences must be genuine differences in *understanding why* empirical phenomena occur. If instead these intellectual differences concern some other kind of understanding, explanationism is left unscathed. Accordingly, an explanationist might argue that differences in EDRs do not genuinely affect understanding-why. Rather, these differences might merely affect our understanding of the *concepts* used to represent explanations, concepts such as Lagrangians, Hamiltonians, connected diagrams, Lorentzian manifolds, etc.[21] If so, conceptualism would have failed to identify a genuine source of intellectual differences.

Conceptualism agrees with part of this objection: in the first instance, reformulating an explanation changes our understanding *of that explanation*. However, non-trivial changes in understanding an explanation entail differences in understanding-why. Conceptualism reframes this claim as a simple bridge principle:[22]

> *Intellectual bridge principle* (IBP): A non-trivial difference in understanding an explanation of p entails a different understanding why p.

According to this bridge principle, organizing the same explanatory information differently can lead to a different understanding-why, as we have seen in the case of Lagrangian and Hamiltonian mechanics. Different ways of understanding an explanation are *non-trivial* provided that they are not merely conventional differences in presenting an explanation. Hence, the intellectual bridge principle excludes a large class of *trivial notational variants* from counting as intellectually significant.[23] For instance, uniformly replacing "5"

[21] I adapt this objection from Khalifa (2017, p. 138), who develops it against Lipton (2009).

[22] de Regt similarly argues that understanding a phenomenon necessarily requires being able to understand a theory (2017, p. 44). However, I disagree with de Regt that understanding a theory is inherently pragmatic or contextual.

[23] Grammatically, "intellectually significant" is analogous to "explanatorily significant." It characterizes differences that matter for understanding—specifically the non-practical, objective dimensions of understanding.

everywhere with "V" in an Arabic numeral system would result in different presentations of many explanations, but these differences would be trivial, rather than intellectually significant. Similarly, recasting an explanation using a left-handed coordinate system rather than a right-handed one would not result in any differences in understanding-why. Although it is difficult to precisely delimit trivial from non-trivial reformulations, my defense of conceptualism only requires clear cases of non-trivial reformulations, such as those developed in Section 3.4. Conceptualism posits that a difference in EDRs is both necessary and sufficient for an intellectually significant difference. Trivial notational variants do not provide different EDRs and hence do not generate intellectual differences.[24]

In response, an explanationist might attempt to reject this bridge principle. However, the IBP follows straightforwardly from the received view of understanding, which explanationism seeks to uphold. Recall that according to the received view, understanding why a phenomenon occurs amounts to grasping an explanation of that phenomenon. Grasping explanations requires that we can represent them, and any way of representing explanations involves concepts. Hence, understanding the relevant explanatory concepts is necessary for understanding-why. Understanding-why is thereby derivative on the way that we have understood this explanation, such as the epistemic dependence relations we have used to present it. Thus, at least some changes in explanatory concepts must lead to concomitant changes in understanding-why. In other words, any account of understanding requires a bridge principle to connect our explanatory concepts with achieving understanding.

With these distinctions in hand, conceptualism straightforwardly identifies the origins of intellectual differences between the equivalent formulations mentioned in Section 3.4. To take one example, the Einstein algebra formalism is markedly different from the standard formulation of general relativity. It teaches us that we don't need to know the standard Lorentzian manifold and metric concepts to provide explanations in general relativity. Instead, we can reorganize all of the relevant explanatory information using algebraic notions, as Geroch (1972) has argued. Since this reformulation changes what we need to know to present explanations, it is not a trivial notational variant of the standard formulation. It thereby satisfies the intellectual bridge principle, leading to a different understanding-why for phenomena explained by general relativity.

[24]Section 2.6.3 defends this criterion for distinguishing trivial from significant reformulations.

By itself, conceptualism does not provide a full-fledged account of scientific understanding. Instead, it illuminates an important facet of understanding that has been neglected in the literature. Due to its minimal commitments, conceptualism can be adjoined with existing accounts of understanding, particularly those allied against explanationism. Although compatible with skills-based accounts of understanding, conceptualism does not assume any special role for skills or abilities. The key insight behind my position is that how a theory-formulation organizes explanatory information matters for understanding. Scientific agents perform no more special a role than grasping this organizational structure. For these reasons, my position is not susceptible to the explanationist strategy against skills-based accounts considered in Section 3.3. Likewise, since conceptualism focuses on how recasting explanations changes understanding, it does not succumb to Khalifa's objections to Lipton's (2009) *understanding without explanation* proposal.

3.6 An Objection against Explanatory Equivalence

In response, an explanationist might reject my argument in Section 3.4 that theoretically equivalent formulations provide the *same* explanation. They might argue that in such cases, one formulation takes explanatory priority. There are at least two candidate sources of explanatory priority. First, one formulation might be physically privileged. For instance, Curiel (2014) privileges Lagrangian mechanics for allegedly encoding the kinematic constraints of classical systems. Second, one formulation might be more fundamental than another. This metaphysical difference would presumably entail a corresponding explanatory difference, wherein the more fundamental formulation provides a better explanation (Sider 2011, p. 61). Differences in joint-carving or perfectly natural properties would then be part of the explanatory nexus. For instance, North (2009) argues that Hamiltonian mechanics is more fundamental than Lagrangian mechanics.[25]

However, this objection sits uneasily within the broader dialectic of Khalifa's explanationism. Recall from Section 3.3 that to uniformly problematize multifarious accounts of understanding, Khalifa adopts a form of explanatory pluralism. Otherwise, we could easily designate some aspects of explanation (e.g. the causal-mechanical ones) as genuinely explanatory while viewing other aspects (such as unification) as mattering for un-

[25]See Section 2.5 for some problems facing appeals to differences in joint-carving or fundamentality.

derstanding but not explanation. Insofar as explanationism requires pluralism, it cannot preclude the interpretation of theoretically equivalent formulations adopted in Section 3.4. It must allow philosophers to interpret cases of theoretically equivalent formulations as being just that: genuinely equivalent both physically and metaphysically.[26] If explanationists instead adopt explanatory monism, they will be unable to systematically recast all purported differences in understanding as explanatory differences. The explanationist is thus caught on the horns of a dilemma. Either they renounce explanatory pluralism and thereby fail to systematically deflate skills-based accounts of understanding, or they maintain pluralism and thereby allow that theoretically equivalent formulations provide the same explanation but different understandings. For those who are happy to reject explanatory pluralism, Section 4.4 poses further problems for explanationism.

3.7 An Extended Illustration: Feynman Diagrams

To illustrate features of understanding that conceptualism illuminates, I consider the case of Feynman diagrams in particle physics. I will argue that Feynman diagrams provide an important way of understanding scattering experiments *at least* in virtue of their formal representational features, setting aside features that might depend on human psychology. I will focus on the formal role that Feynman diagrams perform in representing the connectivity properties of terms in a perturbation expansion.[27] These properties make it unnecessary to consider a vast number of terms in the perturbation expansion. This leads to uniform expressions for simplified formulas that describe particle scattering. Yet as we will see, Feynman diagrams are not unique in providing this representational capacity. When it comes to formally expressing connectivity, other representational frameworks succeed just as well. I leave open whether additional non-psychological, formal features can account for the *prima facie* differences between Feynman diagrams and these other frameworks. My primary goal is to provide a detailed example illustrating my strategy

[26] As Rosenstock et al. note, "it seems far more philosophically interesting to recognize that the world may admit of such different, but equally good, descriptions than to argue about which approach is primary" (2015, pp. 315–16). Chapter 6 further rebuts appeals to differences in joint-carving.

[27] Although it is typically the *diagrams* themselves that are called 'connected' or 'disconnected,' each diagram corresponds to a term in a perturbation series. Hence, we can say that a term is connected if it arises from a connected diagram.

for meeting Khalifa's challenge to accounts of understanding.[28]

The connectivity properties of Feynman diagrams are simply one of their many methodologically beneficial features. Arguably, some of these other features are even more important. Nevertheless, for my goals here, it suffices to identify a single intellectually significant feature, so as to illustrate the kind of epistemic value that conceptualism highlights. Hence, I focus on a single way in which Feynman diagrams contribute non-practical, epistemic value.

Before diving into the gory details for computing scattering amplitudes, it helps to see a schematic overview of the two formulations under discussion. On the one hand, we have what I will call the *naïve approach* to amplitudes. This approach proceeds by naïvely applying perturbation theory, similar to the elementary approach to crystal field theory discussed in Chapter 4. For each scattering process of interest, we can calculate terms in a perturbation expansion to approximate the scattering amplitude. The naïve approach lacks the expressive means to express the connectivity properties of terms. From the standpoint of convenience, this results in having to calculate many more terms in the perturbative expansion. From the standpoint of intellectual significance, failure to express connectivity properties makes it impossible to express the following EDR: only the connected terms in the perturbative expansion contribute non-trivially to the scattering amplitude. The disconnected terms contribute to the trivial part of the scattering matrix, i.e. the identity. This trivial part represents processes where the particles do not interact.

What I will call the *sophisticated approach* to scattering amplitudes remedies this expressive limitation. By expressing the connectivity properties of terms, it provides the aforementioned epistemic dependence relation: in order to calculate the scattering amplitude to a given order in perturbation theory, it suffices to analyze a finite number of connected terms. In other words, it is not necessary to calculate the disconnected terms. This epistemic dependence relation is an instance of modularization. It allows us to decompose the problem of calculating a scattering amplitude into the sub-problems of (i) determining and then (ii) calculating the connected terms. Feynman diagrams are one way to carry out out this sophisticated approach, but other expressive means also provide ways to express connectivity, which is all this approach requires.

[28]For technical background and formal results, see for instance Srednicki (2007, §§8–10) and Lancaster and Blundell (2014, §§16–20, 22, and 24).

These two formulations of scattering amplitudes show how we can have differences in understanding without differences in explanation. Both formulations ultimately provide the same explanation for a given amplitude. They appeal to the same physical laws governing interacting quantum field theories. Likewise, they appeal to the same terms in the perturbation expansion (with the same physical content). Thus, the formulations are compatible rather than competing. The explanation that one provides is (at least) as true (or false) as the other, and for the same physical reasons. Whereas the naïve approach shows by brute force calculation that various terms contribute only trivially, the sophisticated approach takes a shortcut through connectivity. Modularizing the perturbation theory calculation to focus on connected terms does not change the explanation. Rather, it changes how we express this explanation.

Against my claim of explanatory parity, some might object that only the sophisticated approach shows that the disconnected terms are "explanatorily irrelevant," since they do not contribute to the non-trivial, interacting part of the scattering amplitude. Assuming that irrelevant information detracts from an explanation, this would render the naïve approach less explanatory. This objection is misguided for two reasons. First, the disconnected terms are not explanatorily irrelevant: it matters for the scattering amplitude that they do not contribute to the interacting part. If they did contribute, then the interacting part of the amplitude would be different. Second, both approaches do in fact show that the disconnected terms only contribute to the trivial part of the scattering matrix (either by a brute force calculation or by connectivity properties).

Similar remarks apply to the use of selection rules throughout quantum physics and chemistry. It is explanatorily relevant, not irrelevant, that various matrix elements of physical operators vanish. For the purposes of providing an explanation, it does not matter if we determine these vanishing matrix elements by brute force calculations or by elegant symmetry arguments. It does, however, matter for the purposes of understanding the phenomena at hand.

Sections 3.7.1-3.7.3 provide details about computing scattering amplitudes using perturbation theory and Feynman diagrams. Readers who are satisfied with a schematic philosophical understanding of this example can skip ahead to either the philosophical analysis in Section 3.7.4 or to the concluding section of this chapter.

3.7.1 Generating functionals as a Taylor expansion

To compute scattering amplitudes and decay cross-sections for an interacting quantum field theory, one standard method constructs a *generating functional* $Z_I(J)$, where the "I" stands for "interacting" and the "J" stands for a source term for a quantum field. By definition, the generating functional is the probability amplitude for starting out in the vacuum state and remaining in the vacuum state in the presence of a source term J: $Z_I(J) \equiv \langle 0|0\rangle_J$. The point of this subsection is to motivate an expression for the generating functional as a Taylor expansion (Equation 3.5). This then enables the use of perturbation theory to approximate the scattering amplitude.

To construct the generating functional, we start by expressing it as a path integral:

$$Z_I(J) = \int D\phi \, e^{i\int d^4x[L_0+L_I+J\phi]} \tag{3.1}$$

Here, L_0 is the free Lagrangian (devoid of all interaction terms), whereas L_I is the interacting part of the total Lagrangian. To illustrate, we will consider the case of ϕ^3*-theory*, given by the following Lagrangian:

$$L = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 + \frac{1}{6}g\phi^3 \tag{3.2}$$

The free Lagrangian consists of the kinetic and potential energy terms: $L_0 = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2$. The interacting Lagrangian comprises the interaction term: $L_I = \frac{1}{6}g\phi^3$ where g is the coupling strength between the ϕ-fields.

The free field Lagrangian L_0 has a generating functional that we can represent exactly as an exponential of a product of source fields and a free field propagator $\Delta(y-z)$, integrated over the field variables (note that these propagators end up being crucial to determining whether a term is connected or disconnected):

$$Z_0(J) = exp\left[\frac{i}{2}\int d^4x\, d^4x'\, J(x)\Delta(x-x')J(x')\right] \tag{3.3}$$

This lets us express the generating functional for the *interacting* theory as a functional derivative of this free-field generating functional, proportional up to normalization:[29]

$$Z_I(J) \propto exp\left[\frac{i}{6} g \int d^4x \left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^3\right] Z_0(J) \tag{3.4}$$

We can then proceed to apply perturbation theory by expanding $Z_I(J)$ as a Taylor expansion in powers of g and J, which yields the following:

$$Z_I(J) \propto \sum_{V=0}^{\infty} \frac{1}{V!}\left[\frac{ig}{6}\int d^4x \left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^3\right]^V \times \sum_{P=0}^{\infty} \frac{1}{P!}\left[\frac{i}{2}\int d^4y\, d^4z\, J(y)\Delta(y-z)J(z)\right]^P \tag{3.5}$$

Each order of perturbation theory corresponds to the power of g in the corresponding terms, i.e. the value of the expansion parameter V. Hence, to compute the generating functional to first order in perturbation theory, we would compute all terms with $V = 0, 1$. We do this by considering all possible ways that the functional derivatives $\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^3$ can act on the source–propagator–source terms $J(y)\Delta(y-z)J(z)$. Since each functional derivative annihilates a single source term, we see that the remaining number of source terms $J(x)$ is given by $E = 2P - 3V$ (each propagator term supplies $2P$ source terms, and each functional derivative term supplies $3V$ functional derivatives). Clearly, we only require that E is non-negative, so for any given choice of V there will be infinitely many terms corresponding to the summation over P.

3.7.2 Connected vs. disconnected terms and diagrams

The infinite number of terms in Equation 3.5 initially makes the perturbation theory calculation seem intractable: we can seemingly never finish computing all relevant terms to any given order in perturbation theory.[30] There are at least two ways for reducing this infinite number of terms to a computationally tractable finite number of terms. The seemingly most *naïve approach* would proceed by showing that after a certain point, additional propagator terms do not contribute non-trivially to scattering amplitudes and decay rates.

[29]The $\frac{\delta}{\delta J(x)}$ term indicates that we are taking a functional derivative acting on the free-field generating functional $Z_0(J)$.

[30]Note that this infinity of terms has no essential connection to the infinite divergences handled by renormalization methods. I do not consider the complexities of renormalization here. A more in-depth investigation would articulate how formal features of Feynman diagrams also underlie their utility for understanding renormalization.

Physically, these additional propagator terms correspond to particles entering and exiting without scattering. Formally, we would show that once we have fixed V, there is an upper bound on P for the resulting terms to be non-trivial. Ultimately, I will argue that the formal deficiencies of this naïve approach illustrate how Feynman diagrams (and other representational methods considered below) provide a different understanding of scattering processes, while providing no differences *qua* explanation.

A more *sophisticated approach* to handling the Taylor expansion of Equation 3.5 relies on the following insights: for any choice of V, there are a finite number of *connected* terms. Additionally, it is only the connected terms that contribute to the non-trivial, interacting part of the scattering amplitude. Hence, in order to compute the generating functional to a desired order in perturbation theory, we need only consider the corresponding connected terms. This insight lies behind the power and utility of Feynman diagrams. Nevertheless, Feynman diagrams are simply one (particularly perspicuous) method for expressing the connectivity properties of these terms.

At this point, a disclaimer is in order: standard presentations speak only of connected or disconnected *diagrams*, where 'connected diagram' means a path-connected graph. In contrast, the sophisticated approach ascribes the property of connectedness to terms in the perturbation expansion. In general, we can say that a term is connected if and only if it arises from a connected diagram (likewise for a disconnected term). This non-standard usage is necessary in order to make the naïve approach commensurable with the sophisticated approach. It is necessary to focus on properties of the terms in the perturbation expansion, rather than properties of any particular way of representing those terms.

In a position space representation, we can provide a more explicit definition of a connected term. A term is *connected* provided that there is a path of propagators $\Delta(y-z)$ between every pair of remaining source terms $\int d^4x\, J(x)$ and/or "vertex terms" $g\int d^4y$. To illustrate the difference between a connected and a disconnected term, consider a term resulting from the choice $V=2$ and $P=4$:

$$i\int d^4x\, J(x)\, \frac{1}{i}\Delta(x-y)\, ig\int d^4y\, ig\int d^4u\, \left[\frac{1}{i}\Delta(y-u)\right]^2 \frac{1}{i}\Delta(u-v)\, i\int d^4v\, J(v) \tag{3.6}$$

This term is connected because there is a path of propagators connecting the position dummy variables x, y, u, and v. For instance, the source term $i\int d^4x\, J(x)$ is connected to the vertex term $ig\int d^4y$ by the propagator $\frac{1}{i}\Delta(x-y)$. Similarly for the source term $i\int d^4v\, J(v)$, which is connected to the vertex term $ig\int d^4u$ by the propagator $\frac{1}{i}\Delta(u-v)$. In contrast, consider the following term that results from the choice $V=2$ and $P=5$:

$$i\int d^4x\, J(x)\,\frac{1}{i}\Delta(x-y)\, ig\int d^4y\, ig\int d^4u\,\left[\frac{1}{i}\Delta(y-u)\right]^2\frac{1}{i}\Delta(u-v)\, i\int d^4v\, J(v) \\ \times\, i\int d^4w\, J(w)\,\frac{1}{i}\Delta(w-z)\, i\int d^4z\, J(z) \tag{3.7}$$

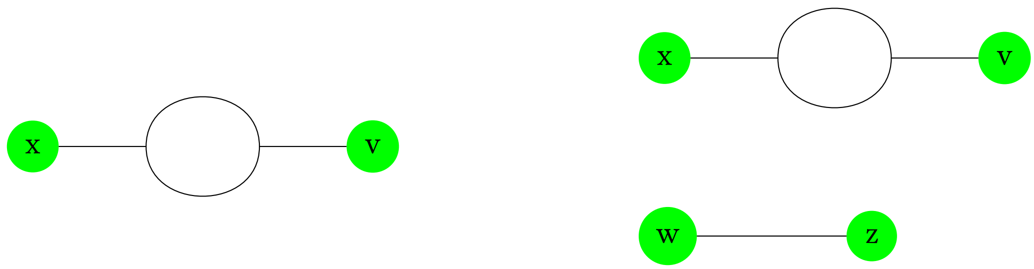
This term is disconnected because there is no propagator connecting the source term $i\int d^4x\, J(x)$ defined at x to the source term $i\int d^4w\, J(w)$ defined at w. Instead, this term is a product of two connected terms.

Exploiting the connectivity properties of terms greatly simplifies our representation of the interacting generating functional $Z_I(J)$. As shown by the Taylor expansion in Equation 3.5, $Z_I(J)$ is proportional to the sum of all terms. However, using the language of connected terms, we can prove that *the generating functional is proportional to the exponential of the sum of all* connected *terms* (Srednicki 2007, p. 65). Hence, to compute the generating functional to any order in perturbation theory, we need only focus on connected terms, rather than disconnected terms as well. This illustrates two distinct ways of understanding the generating functional, both of which are physically equivalent: either (1) as the sum of all terms or (2) as the exponential of the sum of connected terms. For convenience, I'll continue to refer to the first method as the *naïve approach* to computing the generating functional. Likewise, I'll call the second method—which uses connectivity—the *sophisticated approach*.

3.7.3 Feynman diagrams and Feynman rules

In the position space representation that we have been working in, the integral terms themselves express their connectivity properties. Thus, position space integral terms provide one way of implementing the sophisticated approach. Feynman diagrams provide another method for implementing this approach: they also express the connectivity properties of terms in the Taylor expansion of the generating functional. Hence, from a formal standpoint, the corresponding position space Feynman diagrams do not add anything, at least when it comes to connectivity. We will see shortly that this is decidedly not the case when we move to a momentum space representation.

To see how the position space Feynman diagrams express exactly the same information as the corresponding integral terms, consider again the terms in Equations 3.6 and 3.7. These terms correspond to the position space Feynman diagrams in Figure 1.



(a) Connected Diagram

(b) Disconnected Diagram

Figure 1: Two position space Feynman diagrams

These diagrams follow from the corresponding *position space Feynman rules* for ϕ^3-theory, which establish correspondences between elements of a diagram and integral terms:

$$\text{Source terms (external blobs):} \quad i\int d^4x\, J(x) \tag{3.8}$$

$$\text{Propagators (internal lines):} \quad \frac{1}{i}\Delta(x-y) \tag{3.9}$$

$$\text{Vertex terms (intersection point of three lines):} \quad ig\int d^4x \tag{3.10}$$

However, when it comes to computing scattering cross-sections, it is often much more computationally convenient and tractable to work in a momentum space representation. Moving to momentum space eliminates all source terms (the external blobs) and replaces position dummy variables with momentum dummy variables $\{k_1, k_2, \ldots\}$, one for each line. Eliminating the source terms leads to a simpler representation for the external lines.

The *momentum space Feynman rules* for ϕ^3-theory are the following:[31]

$$\text{External lines:} \quad \text{a factor of } 1 \tag{3.11}$$

$$\text{Propagators (internal lines) with momentum } k\text{:} \quad \frac{-i}{k^2+m^2-i\varepsilon} \tag{3.12}$$

$$\text{(Unconstrained) internal loop propagator:} \quad \int \frac{d^4k}{(2\pi)^4} \tag{3.13}$$

$$\text{Vertex terms (intersection point of three lines):} \quad ig \tag{3.14}$$

Here, m corresponds to the mass of the particle (the lowest order excitation of the ϕ-field). The factor of $-i\varepsilon$ functions to keep integrals well-defined by analytically continuing into the complex plane.

From these rules, we can immediately see that the integral terms corresponding to momentum space Feynman diagrams are unable to independently express connectivity properties, while the diagrams themselves can. Considering again our earlier examples of the connected and disconnected terms in Equations 3.6 and 3.7, Figure 2 depicts the corresponding momentum space Feynman diagrams.

(a) Connected Diagram

(b) Disconnected Diagram

Figure 2: Two momentum space Feynman diagrams

These diagrams clearly maintain the connectivity properties of their corresponding position space terms: the former is connected, whereas the latter is disconnected.

In contrast, when we apply the momentum space Feynman rules, we see that the corresponding momentum space integral terms for *both* diagrams equal the following:

$$(ig)^2 \int \frac{d^4\ell}{(2\pi)^4} \frac{-i}{\ell^2+m^2-i\varepsilon} \frac{-i}{(k_1-\ell)^2+m^2-i\varepsilon} \tag{3.15}$$

[31]Note that since momentum is conserved at each vertex, some values of momentum are constrained in terms of the momenta on adjoining lines.

Hence, we see that these two rather distinct diagrams both correspond to the same integral term in the momentum space representation. This is a consequence of the fact that external lines in the momentum space representation correspond to a factor of one.[32]

3.7.4 Differences in understanding

We can now examine in more detail how Feynman diagrams illustrate my conceptualist account of scientific understanding. Putting together the pieces, both the naïve approach and the sophisticated approach provide two methodologically distinct but physically equivalent ways of explaining scattering amplitudes. First, consider the naïve approach. From the Taylor expansion for the generating functional, we compute every nontrivial position space term to our desired order in perturbation theory. This includes both connected and disconnected terms. Next, we re-express these terms in a momentum space representation and compute the generating functional. From this generating functional, we can compute the desired scattering amplitudes.

Second, consider the sophisticated approach. Starting from the same Taylor expansion, we exploit the connectivity properties of the terms, enabling us to neglect all disconnected terms. We ultimately find that the generating functional is given by the exponential of the sum of connected terms (equivalently diagrams). Thus, we determine all connected momentum space diagrams (up to our desired order in perturbation theory), apply the Feynman rules, and arrive at the corresponding momentum space integral terms. We then compute the generating functional on the basis of these.

A few remarks demonstrate that these two formulations are on a par *qua* explanations of scattering amplitudes. They both appeal to identical sets of physical laws (the laws underlying interacting quantum field theories). They have identical physical ontologies, specified by the ϕ^3-theory Lagrangian and whatever preferred (partial) interpretation of quantum field theory we might have. They both also lead to identical generating functionals and thus identical scattering amplitudes. It is not the case that one formulation makes additional physical or mathematical assumptions that the other does not. Furthermore, they have the same explanatory scope: if one formulation works for explaining a

[32]Technically, we could distinguish these two diagrams by taking into account the conservation of momentum of incoming and outgoing momentum variables. For simplicity, I neglect this subtlety. A slightly more complicated example would show more rigorously that momentum space integral terms cannot express connectivity.

given scattering amplitude then the other will work as well. There is also seemingly no possible causal difference between the explanations they provide: they both appeal to the same physical processes. Thus, we have good reason to believe that the naïve and the sophisticated approaches represent the same explanation.

Nevertheless, we have already seen that these two reformulations have stark conceptual differences, and it is these differences that my account of understanding captures. Only the sophisticated approach succeeds at expressing the generating functional as the exponential of a sum of connected momentum space terms/diagrams. This difference stems from the fact that although each momentum space Feynman diagram corresponds to a momentum space integral term via the Feynman rules, the diagram but not the integral encodes connectivity properties. This shows that the momentum space Feynman diagrams underwrite an epistemic dependence relation that the corresponding integral terms do not: the diagrams can express connectivity properties. Recognizing the relationship between connectivity and the generating functional modularizes the scattering amplitude problem into the two sub-problems of (i) determining the connected terms and (ii) calculating their contribution. Hence, these diagrams provide a way of understanding scattering amplitudes that the momentum space integral terms cannot provide.

Besides modularizing the problem, the momentum space Feynman diagrams also provide a *uniform treatment* of generating functionals in terms of connected diagrams. In this context, uniformity amounts to the fact that the diagrams—or, more generally, any representation that succeeds at expressing connectivity—enable us to say *for all scattering processes* that we need not compute disconnected terms. Hence, it is not just that for any *particular* scattering process we end up with the generating functional as an exponential of the sum of connected terms: we know that this result holds even before we start computing. In contrast, by failing to exploit connectivity, the naïve approach lacks the capacity to uniformly express this formal relationship. The closest it could come to expressing this relationship would be to verify it in *particular* cases. For instance, after arriving at our generating functional via the naïve approach, we could subsequently show that this generating functional is equal to the exponential of the sum of all of the "connected" terms. I place "connected" in scare quotes here because we are assuming that the naïve approach is formally unable to express connectivity.

More precisely, in this scenario we would be showing the following: the exponential of the sum of all momentum integral terms that in fact correspond to connected diagrams yields the same generating functional. Yet, by the lights of the naïve approach, this equivalence might as well be a coincidence. Note that this is not a failure of explanatory scope: as noted above, both formulations succeed for the same set of cases. Instead, it amounts to an expressive limitation. It leads to a difference in the way that the generating functional is understood.[33]

Lange's (2017) account of distinctively mathematical explanations leverages this notion of coincidence to great effect. His account could thereby capture the difference in uniformity between the naïve and sophisticated approaches. However, it requires that the world have a sufficiently rich modal structure. In Section 4.4.2, I argue that this ontological requirement generates an underdetermination problem, similar to that facing fundamentalism (see Section 2.5). Here, all I have assumed is that we have an epistemic difference between seeming-coincidences and seeming-non-coincidences. For all I have said, this relationship between generating functionals and connected terms could still be a coincidence. Either way, we learn the following: any generating functional can be expressed in terms of its connected terms. The issue of ontic coincidences aside, the naïve approach cannot even say this much.

To summarize then, it is not the case that we *need* Feynman diagrams to compute scattering amplitudes. Rather, Feynman diagrams provide a particularly convenient way of expressing the connectivity properties of terms in the perturbation expansion. It is these connectivity properties that enable us to simplify the generating functional, making it unnecessary to compute a large number of terms. However, at least when it comes to these connectivity properties, the position space integral representation succeeds just as well. Furthermore, there are additional ways to represent these connectivity properties, such as through Wick contractions of vacuum expectation values. This approach is used, for example, in the canonical quantization approach to computing scattering amplitudes for interacting field theories.[34] Ultimately though, the expressive differences between mo-

[33]One might object that only the sophisticated approach succeeds at *explaining* why the generating functional equals the exponential of the sum of connected terms. Nevertheless, explaining this epistemic dependence relation is a separate *explanans* than explaining the scattering amplitude. In this case study, we are focused only on explaining scattering amplitudes. Additionally, if explanation is ontic, then plausibly the naïve approach can explain this EDR just as well, simply by representing the relevant states of affairs involved.

[34]See, for instance, Lancaster and Blundell (2014), §§18–19.

mentum space Feynman diagrams and momentum space integral terms shows that these diagrams provide an important way of understanding the generating functional that the momentum space integral terms cannot provide, despite each diagram corresponding to an integral term via the momentum space Feynman rules. This difference illustrates how the formal language we choose can change how we understand the physical quantities we compute.

3.8 Conclusion

I have argued that theoretically equivalent formulations provide a clear counterexample to explanationism. Whereas explanationism holds that all intellectual differences arise from explanatory differences, equivalent formulations show that some differences in understanding-why do not reduce to explanatory differences. To accommodate these intellectual differences, I have defended *conceptualism*. Conceptualism argues that understanding-why involves not only the explanatory content that we have understood, but also the way that we have understood it. In particular, it claims that equivalent formulations manifest intellectual differences whenever they provide different *epistemic dependence relations*. These are differences in what we need to know or what suffices to know to solve scientific problems. By characterizing how reformulations change understanding, conceptualism addresses complementary lacunae in current accounts of both scientific understanding and theoretical equivalence. In this way, conceptualism supplements existing anti-explanationist accounts of scientific understanding. By adopting conceptualism, these accounts can forestall the challenge from explanationism and illuminate understanding beyond scientific explanation.

To illustrate my strategy, I considered some simple Feynman diagrams in particle physics. I argued that Feynman diagrams provide a distinct way of understanding scattering experiments when compared to the momentum space representation of the corresponding integrals. Feynman diagrams provide this understanding by expressing the connectivity properties of terms in a Taylor expansion for the generating functional. Although the corresponding momentum space integral terms can perform the same computational roles as the Feynman diagrams, they cannot express these connectivity properties. This prevents them from modularizing the scattering amplitude problem and uniformly expressing the generating functional in terms of connected terms. When we use the momentum space integral terms in a naïve approach to explaining scattering amplitudes, we miss out on a key epistemic dependence relation. In contrast, using a sophisticated approach (such as momentum space Feynman diagrams) exploits this EDR to modularize and uniformly express the scattering amplitude problem, providing a different understanding. However, both these naïve and sophisticated approaches are physically equivalent and share identical explanatory scope: they lead to identical generating functionals from identical assumptions. They therefore provide a difference in understanding with no difference in explanation.

Chapter 4:

Reformulating through Symmetry

4.1 Introduction

Chapter 3 used cases of theoretical equivalence to undermine explanationism. Such cases show that there are features of scientific understanding that go beyond scientific explanation. Nevertheless, one might question the scope and force of this argument. Does it only apply in cases of theoretically equivalent reformulations? If so, that would be a severe limitation, since many compatible reformulations are not theoretically equivalent. This includes many symmetry arguments in physics and chemistry, exemplified by the case study I analyze in this chapter.[1]

This chapter applies conceptualism to one of the simplest yet sufficiently rich examples of a general kind of symmetry argument in physics and chemistry. The example comes from an idealized model known as crystal field theory. Section 4.2 introduces three compatible formulations of crystal field theory: the elementary, non-group-theoretic, and group-theoretic approaches. Each approach references the same ontic explanatory features, while nonetheless leading to different understandings of the phenomena. I will argue that these formulations illustrate another kind of counterexample to explanationism. Other examples of symmetry-based reformulations include the Wigner–Eckart matrix element theorem (Hunt 2021a), selection rules in spectroscopy (Hunt 2014), and symmetry-based explanations of hydrogen's energy spectrum (Singer 2005).

Since symmetry arguments are not theoretically equivalent to the elementary methods they reformulate, we need a new argument against explanationism in this context. Otherwise, an explanationist could argue that explanatory differences ground the relevant intellectual differences. They might argue that in cases of symmetry-based reformulation, all differences in understanding arise from explanatory differences. If so, this would

[1] Sections 4.2-4.4 and the appendix of this chapter were published in Hunt (forthcoming).

severely limit the scope of conceptualism. It would show that in many cases, conceptualism is at best redundant, adding nothing to existing accounts of scientific explanation. In this chapter, I rebut this more general challenge from explanationism. In doing so, I extend the scope of conceptualism as an account of understanding that goes beyond explanatory differences. Both the elementary and the symmetry-based approaches provide different understandings of the same phenomena, despite describing the same ontic explanatory information. Section 4.2 illustrates this moral in detail. Moreover, my argument in this chapter does not presuppose that categorical equivalence is a good standard for theoretical equivalence.

My account of how symmetry arguments contribute to understanding involves disentangling three compatible formulations. First, there are *elementary approaches*, which proceed on a case-by-case basis without appealing to symmetry. Often in physics and chemistry, elementary approaches involve a brute-force application of perturbation theory to each system of interest. Second, in *non-group-theoretic approaches*, we make the system's symmetries explicit but without using an abstract language for symmetry. Finally, *group-theoretic approaches* take advantage of symmetry by using the more sophisticated mathematics of group representation theory. Section 4.2.3 shows how at each stage in this process of reformulation, we acquire different epistemic dependence relations, leading to different understandings of the phenomena.

In particular, symmetry-based reformulations affect understanding through two general kinds of epistemic dependence relations: modularization and unification. *Modularization* occurs when a formulation breaks a problem or a why-question into separately treatable sub-problems. Modularizing a problem shows that some parts of it can be treated independently of other parts. *Unification* occurs when a single derivation and its solution applies to a family of different systems that all display shared behavior. For instance, organizing systems into symmetry-based families unifies them. Noticing such epistemic dependence relations changes our understanding of a given phenomenon by clarifying what suffices or is necessary to understand it. Furthermore, this kind of intellectual difference does not rely on any particular details about agents, skills, or capacities. We can thereby abstract away agents, analyzing EDRs as objective, agent-independent features of a theory formulation.[2]

[2]This non-agentive approach to understanding captures one facet of understanding that does not de-

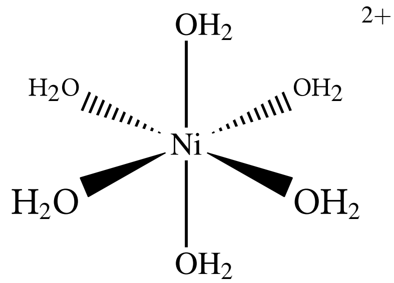
An explanationist might wonder whether current accounts of explanation can accommodate these intellectual differences. Doing so would render conceptualism redundant. Sections 4.3 and 4.4 consider and rebut two objections against conceptualism, stemming from this worry. The first objection argues that existing accounts of explanation can easily accommodate the intellectual differences I identify. Considering two leading accounts of causal explanation, I show that this is not the case. The second objection argues that what I am calling cases of the same explanation are in fact different explanations and can be accommodated as such. I argue that this response faces a skeptical challenge that conceptualism avoids.

Finally, I consider what it takes for one formulation to provide a *better understanding* than another. Section 4.5 proposes an expressivist account of better understanding. I argue that comparative judgments of understanding express preferences for intellectual features or properties of arguments. Judging one formulation to be intellectually better than another is to express a non-cognitive attitude of *being for intellectually preferring* that formulation. Expressing these preferences amounts to expressing acceptance of a system of norms. Section 4.6 considers a number of different norms that might govern intellectual preferences. Accepting or rejecting these norms leads to particular first-order claims about comparative understanding, such as the claim that the group theoretic approach provides a better understanding of crystal field theory than the elementary approach.

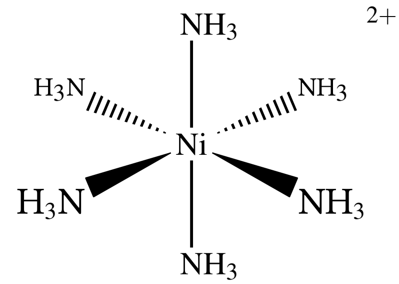
4.2 A Case Study from Crystal Field Theory

Crystal field theory provides an idealized model for describing properties of coordination complexes. These consist of a positively charged metal ion surrounded by negatively charged or polarized species known as 'ligands.' Figure 3 shows two examples: nickel(II) hexahydrate and nickel(II) hexammine. Both complexes comprise a Ni^{2+} ion bound to six ligands occupying the vertices of an octahedron (see Figure 6). Often, the color of coordination complexes changes according to the ligands bound to the metal ion. Whereas nickel(II) hexahydrate is green, nickel(II) hexammine is purple. Chemists use crystal field theory to understand these differences in color, along with differences in thermodynamic and magnetic properties (Figgis and Hitchman 2000).

pend on agents, allowing agent-dependent features to be subsequently added. Ultimately, psychological 'grasping' should be treated naturalistically, sensitive to the concerns of Trout (2007).



(a) Nickel(II) Hexahydrate. $[Ni(H_2O)_6]^{2+}$



(b) Nickel(II) Hexammine. $[Ni(NH_3)_6]^{2+}$

Figure 3: Octahedral coordination complexes

To explain these properties, chemists focus on how the valence electrons of the metal ion change when surrounded by ligands. For instance, an isolated Ni^{2+} ion has eight valence electrons that occupy five energetically 'degenerate' orbitals, meaning that they have the same energy (depicted by the left side of Figure 4).[3] Surrounding Ni^{2+} with ligands breaks this degeneracy, causing previously degenerate orbitals to 'split' into new energy levels with new degeneracies. Crystal field theory describes this splitting phenomenon by treating ligands as point dipoles that create an electrostatic 'crystal' field, perturbing the energy levels of the metal ion.[4] In the case of nickel(II) hexahydrate, the fivefold degenerate valence orbitals split into two new levels that are two-fold and three-fold degenerate, shown in Figure 4. The difference between these energy levels is denoted 'Δ_O.' Electronic transitions between these levels help explain the characteristic colors of many metal complexes.

Crystal field theory solves three connected problems about electronic structure, each posing its own why-question. The 'splitting problem' is to determine how many new energy levels form from a previously degenerate energy level. The 'degeneracy problem' is to determine how many orbitals constitute each new level, i.e. its degeneracy. Finally, chemists estimate the energy difference Δ_O by finding the eigenvalues of each new energy level, giving rise to the 'eigenvalue problem.' For brevity, I will refer to these three problems collectively as *the crystal field theory problem*. Section 4.2.1 begins by sketching three different compatible approaches to explaining this phenomenon. Armed with

[3] Atomic orbitals are one-electron wavefunctions used to approximate the overall state of an atom or molecule.

[4] Crystal field theory idealizes interactions as purely ionic, neglecting chemical bonds between the metal ion and ligands. It underlies more sophisticated models such as ligand field theory (Cotton 1990, p. 254). For simplicity, I suppress additional philosophical issues pertaining to idealization, since the same questions about reformulations arise outside this context.

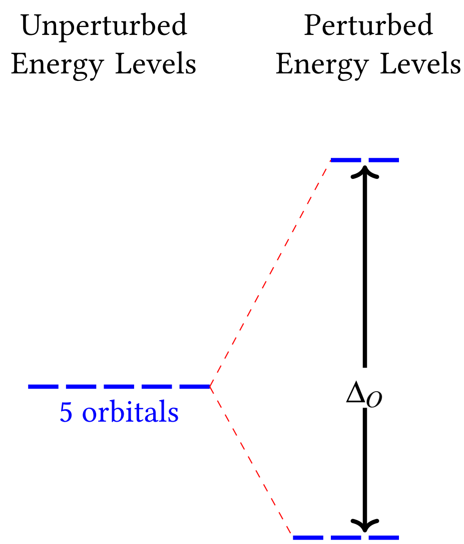


Figure 4: Splitting of valence orbitals in an octahedral crystal field

these approaches, Section 4.2.2 develops them as a counterexample to explanationism. They provide different ways of understanding crystal field theory without concomitant explanatory differences. Finally, in Section 4.2.3, I show how conceptualism easily accommodates the intellectually significant features of this case study. The group-theoretic approach modularizes and unifies crystal field theory by providing distinctive epistemic dependence relations. Throughout, I will focus on nickel(II) hexahydrate as a concrete example, although my discussion applies more generally.

4.2.1 Three approaches to crystal field theory

The first approach to crystal field theory is 'elementary' in the sense that it makes no explicit appeal to symmetry properties of the molecule. Instead, it relies entirely on perturbation theory, approximating the eigenvalues of the coordination complex relative to those of the unperturbed, free metal ion. We begin by measuring the electrostatic potential, representing it as a perturbation operator H'. The eigenvalues of this perturbation operator provide a first-order correction to the known energy states of the free metal ion. We calculate these eigenvalues by solving a 'secular equation' (Equation A.1), which functions as the relevant law-like statement for this explanation. With the eigenvalues in hand, the splitting and degeneracy follow immediately. The number of distinct eigenvalues and their degeneracies corresponds to the number of new energy levels and their degeneracies. For nickel(II) hexahydrate, we obtain two distinct eigenvalues that are three-fold and two-fold degenerate. Figure 5a represents the schematic structure of this approach.[5]

The second approach relies on the same schematic structure: it uses perturbation theory to calculate the eigenvalues, from which the energy-level structure follows. However, we now take explicit advantage of symmetry, although without using the formal apparatus of group representation theory. Hence, I will refer to this first symmetry-based formulation as the *non-group-theoretic approach*.[6] Unlike the elementary approach, we begin by characterizing the electrostatic potential in terms of the symmetry of the coordination complex. For nickel(II) hexahydrate, the resulting potential (Equation A.2) applies to any coordination complex with six ligands at the vertices of an octahedron. We then follow the same procedure as the elementary approach but now using this symmetry-based potential. Solving the secular equation leads to two distinct eigenvalues: $\lambda_1 = -\frac{2}{5}\Delta_O$ (three-fold degenerate) and $\lambda_2 = \frac{3}{5}\Delta_O$ (two-fold degenerate), expressed in terms of their energy difference Δ_O. As in the elementary approach, the splitting and degeneracy follow immediately from these eigenvalues. The two distinct eigenvalues and their degeneracies entail that two new energy levels form that are three-fold and two-fold degenerate.

In the third approach, we take advantage of not only symmetry but also the formal apparatus of group (representation) theory. This *group-theoretic approach* extensively reformulates the crystal field theory problem, leading to a dramatically different organizational structure, shown in Figure 5b.[7] Rather than deduce the splitting and degeneracy from the eigenvalues (as in the other two approaches), we now determine them without solving a secular equation. To begin, we identify the symmetry groups of both the free metal ion and the coordination complex. An unperturbed metal ion, such as Ni^{2+}, is invariant under arbitrary rotations, so its symmetry group is the rotation group. In the case of nickel(II) hexahydrate, since its ligands sit at the vertices of an octahedron, its symmetry group is accordingly the octahedral group.

The next step is to extract information about the energy levels from these symmetry groups, using the mathematics of group representations. Recall from Figure 4 that our task is to determine how the initially five-fold degenerate valence orbitals of Ni^{2+}

[5] See Appendix A for a more detailed account of each formulation.

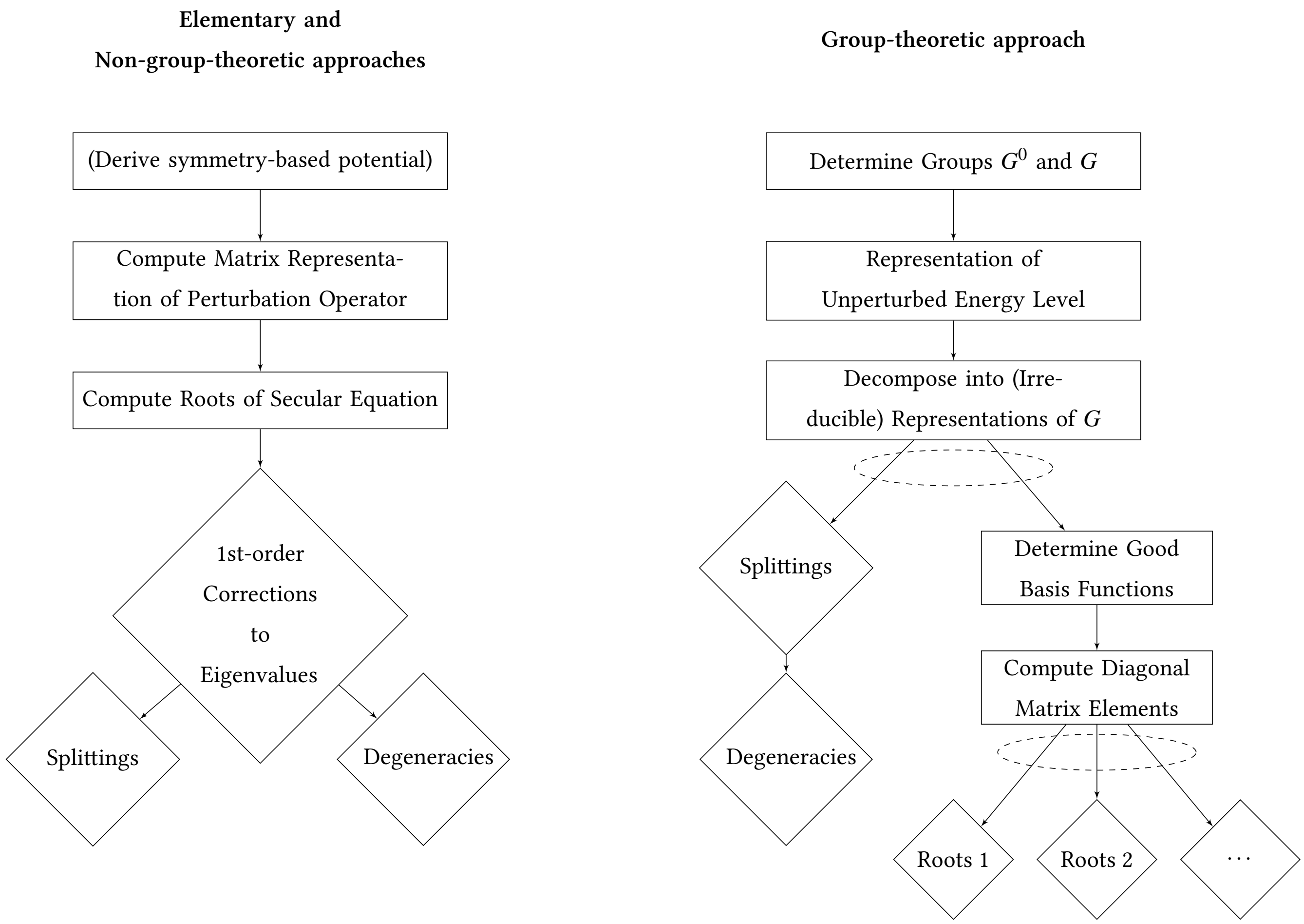
[6] See Dunn et al. (1965, pp. 9–16) and Figgis and Hitchman (2000, pp. 30–38) for detailed applications of this approach.

[7] Adopting an abbreviation common throughout physics and chemistry, 'group theory' will typically refer more precisely to 'group representation theory.' See Appendix A.3 for details.

rearrange into the new energy-level structure of the coordination complex. For both the initial and final systems, each distinct energy level corresponds to a 'representation' Γ of the corresponding symmetry group. Hence, to determine the new splitting and degeneracy, it suffices to determine how many representations of the octahedral group occur (corresponding to the number of new energy levels) and their dimensions (corresponding to the degeneracy of each energy level). First, we determine the representation of the rotation group associated with the nickel ion's valence orbitals. We then exploit a precise mathematical relationship characterizing how this initial representation from the rotation group 'decomposes' into a sum of new representations from the octahedral group. Executing a simple algorithm (demonstrated in Appendix A.3), we find that the initial representation decomposes into two new representations of dimensions two and three. This solves the splitting and degeneracy problems: two new energy levels form that are, respectively, two-fold and three-fold degenerate.

Finally—as in the other two approaches—the group-theoretic approach uses perturbation theory to solve the eigenvalue problem. The key difference is that group theory reorganizes the secular equation using properties of the representations. By knowing the group representations of the new energy levels, we can diagonalize the perturbation operator, H'. Diagonalization provides a separate secular equation for each distinct energy level. This modularizes the eigenvalue problem into a separate sub-problem for each distinct eigenspace. We learn that we can calculate each distinct eigenvalue separately, rather than solving a larger secular equation for all of them.

To summarize central intellectual differences between the three approaches, we can represent the structure of their solution procedures as flowcharts. I represent the elementary and non-group-theoretic approaches together, since they differ only in the first step, namely whether or not we first construct a symmetry-based form for the electrostatic potential. Figure 5 shows how the additional epistemic dependence relations provided by group theory restructure the solution procedure. The dashed ovals indicate modularization, where we have broken a problem into separately treatable sub-problems. Group theory shows us how to separate the splitting and degeneracy problems from the eigenvalue problem, indicated by the first dashed oval in Figure 5. Furthermore, group theory separates the eigenvalue problem into a separate problem for each distinct eigenvalue, indicated by the second dashed oval. Finally, the flowchart indicates that symmetry properties are sufficient for determining the splitting and degeneracy, illustrating how group theory unifies the crystal field theory problem. I expand on these points in Section 4.2.3, but first I clarify how the three approaches pose a problem for explanationism.



(a) The parenthetical first step applies only to the non-group-theoretic approach.

(b) The dashed ovals indicate modularization. See the appendix for a description of irreducible representations.

Figure 5: Schematic step-by-step flowcharts for the three approaches

4.2.2 A problem for explanationism

With the approaches to crystal field theory before us, I will now show how they pose a serious challenge to explanationism. To simplify the exposition, I will focus on how the two symmetry-based approaches explain the splitting and degeneracy (i.e. the energy-level structure). Similar points arise when comparing these two approaches to the elementary approach, along with considering how each approach explains the eigenvalues. My argument involves establishing three premises, which together entail that explanationism provides an incomplete account of the objective and non-pragmatic (or 'non-practical') dimensions of understanding. First, I will show that—on many accounts of explanation—both symmetry-based approaches provide not only derivations but also explanations of the relevant phenomena.[8] Second, I will argue that both approaches reference the same ontic explanatory information; hence, they do not involve explanatory differences. Finally, I will argue that the approaches nevertheless provide different understandings of the crystal field theory phenomena. Hence, explanationism is incorrect: not all intellectual differences stem from corresponding explanatory differences.

Since explanationism is concerned only with explanatory understanding (i.e. understanding-why), my case study poses a problem only if the compatible formulations are genuine explanations, rather than mere derivations. However, due to persistent disagreements about the nature of explanation, it is impossible to conclusively demonstrate that these approaches are explanatory.[9] The best one can do is motivate interpreting them as such. To this end, it suffices to note that many established accounts of explanation would treat these three approaches as explanatory.[10] To make just one representative analysis, I will consider Woodward and Hitchcock's (2003a) *manipulationist* (or 'interventionist') account of causal explanation, since it avoids many well-known problems facing 20th-century accounts of explanation.

Manipulationism recasts explanation as the pursuit of answering *what-if-things-had-been-different questions*: how would the explanandum have differed if one of the explanans had been changed? For instance, we might wonder how a coordination complex's energy levels would have differed if we had changed its symmetry. Answering these what-if questions requires an *explanatory generalization*: a law-like statement characterizing how the explanandum depends on the explanans, in function–variable form. This lets us derive the explanandum from input variables characterizing the explanans (such as initial or boundary conditions). To qualify as an explanatory generalization, a law-like statement must be invariant under 'testing interventions' (Woodward and Hitchcock 2003b, p. 182). This

[8]Section 4.4 extends my argument to explanationists who might refuse to grant this premise.

[9]As Section 2.7 points out, the controversial nature of scientific explanation provides another reason to prefer conceptualism over explanationism when it comes to making sense of reformulations.

[10]These derivations also count as explanatory on Hempel and Oppenheim's (1965 [1948]) deductive-nomological model, Railton's (1981) ideal explanatory text account, Lewis's (1986, pp. 217–21) similar account of causal explanation, and Kitcher's (1989) unificationist account. I briefly discuss the latter in Section 4.3.2.

means that the generalization must continue to hold even as we intervene on the system of interest, changing its explanans variables within some nontrivial range.[11] Coulomb's law provides a paradigmatic example, characterizing how the electrostatic force between two charged bodies depends on their charges and the distance between them.

Manipulationism straightforwardly renders both symmetry-based approaches—and also the elementary approach—as genuinely explanatory. In each formulation, the secular equation (A.1) functions as a suitable explanatory generalization for solving the eigenvalue problem. The secular equation remains invariant under a wide range of interventions, including modifying the charges and configuration of the surrounding ligands. It thereby answers many kinds of what-if-things-had-been-different questions. In the non-group-theoretic approach, we explain the splitting and degeneracy as a consequence of these eigenvalues. Thus, by causally explaining the eigenvalues, we also causally explain the energy-level structure. In the group-theoretic approach, a different explanatory generalization—the character decomposition formula—characterizes how the splitting and degeneracy depend on the symmetry.[12] Since the character decomposition formula is also invariant under a variety of interventions, manipulationism would interpret this as a genuine explanation too. Hence, both approaches provide not merely derivations but also causal explanations of the splitting and degeneracy.

Next, we must show that both symmetry-based approaches reference the same ontic explanatory information. Otherwise, an explanationist could seek to reduce any intellectual differences between the approaches to concomitant explanatory differences. In both cases, we appeal to the geometric arrangement of the coordination complex, i.e. its symmetry, to determine its energy-level structure. The non-group-theoretic approach uses this structural information to determine the symmetry-based form of the potential (Equation A.2). This potential is then fed into the secular equation to determine the quantitative form of the eigenvalues. Although the symmetry-based potential technically references the charge of the central metal ion, this information could be suppressed without changing the derivation of the energy-level structure. Likewise, in the group-theoretic approach, we appeal to the geometric structure of the coordination complex to determine its ab-

[11]These testing interventions need not be experimentally feasible or even physically possible (although they often are); an intervention simply needs to be "logically or conceptually possible" (Woodward 2003, p. 132).

[12]Appendix A.3 describes this formula (Equation A.5) in detail.

stract symmetry group. Then, using mathematical properties of this abstract symmetry group—namely, properties of its representations—we determine the energy-level structure. Hence, in both cases, we appeal to the same ontic explanation, namely the same state of affairs in the world.[13]

Finally, it remains to show that the two approaches provide different ways of understanding the crystal field theory phenomena, despite relying on the same ontic explanatory information. In the non-group-theoretic approach, we understand the splitting and degeneracy as a consequence (or feature) of the eigenvalues. In contrast, the group-theoretic approach provides a way of understanding the splitting and degeneracy as a consequence of symmetry independently of the quantitative form of the eigenvalues. Below, I will describe how conceptualism accommodates this intellectual difference in terms of 'modularization,' a property that some epistemic dependence relations possess. Moreover, the two approaches differ in how they unify coordination complexes into symmetry-based families. In the non-group-theoretic approach, we understand the system's features as an instance of a particular instantiation of octahedral symmetry. Whereas in the group-theoretic approach, our understanding does not depend on the particular instantiation of octahedral symmetry; it applies to any possible instantiation of this symmetry. These differences in unification amount to differences in understanding.[14]

To summarize, the two symmetry-based approaches to crystal field theory provide a counterexample to explanationism. They each explain the splitting and degeneracy while referencing the same ontic explanatory information. Nevertheless, they provide objective and non-practical differences in understanding why the phenomenon occurs. Hence, they show that explanationism is incomplete: it fails to account for all relevant intellectual differences. In the next section, I will demonstrate how conceptualism easily accommodates the intellectual differences between these two approaches, interpreting them in terms of different epistemic dependence relations.

---

[13]Similarly, when it comes to explaining the eigenvalues, each of the three approaches references the same explanatory information, including the charge of the central metal ion, the charges of the ligands, the arrangement of the ligands and metal ion, and the secular equation. This provides a further counterexample to explanationism.

[14]Many philosophers nevertheless view unification as having no bearing on explanation (see Section 4.3).

4.2.3 Illustrating conceptualism

Unlike explanationism, conceptualism accounts for the intellectual differences between compatible formulations. As an illustration, I will show how conceptualism interprets the differences between the two symmetry-based approaches. These amount to organizational differences of the same ontic explanatory information, leading to objective and non-practical differences in understanding. First, I will explain how the notion of modularization accommodates two key intellectual differences between the approaches. Second, I will explain how unification accommodates a remaining intellectual difference.

When it comes to understanding the splitting and degeneracy, the two symmetry-based approaches provide different understandings because they appeal to different EDRs. In the non-group-theoretic approach, we rely on the EDR that knowledge of the quantitative form of the eigenvalues is sufficient for knowledge of the splitting and degeneracy. In contrast, the group-theoretic approach modularizes the crystal field theory problem into distinct sub-problems, solving for the splitting and degeneracy without solving for the eigenvalues. This modularization constitutes the following EDR: knowledge of the eigenvalues is not necessary for knowledge of the splitting and degeneracy. Instead, it is possible to derive the splitting and degeneracy without knowing even the quantitative form of the eigenvalues. These different EDRs thereby provide different ways of understanding the splitting and degeneracy, in virtue of how they structure the derivation. A similar moral about modularization applies to the eigenvalue problem. In the non-group-theoretic approach, we understand the eigenvalues from a single secular equation, whereas in the group-theoretic approach, we modularize the larger secular equation into a set of smaller secular equations, one for each distinct eigenspace. Thus, we learn that we can understand each distinct eigenspace separately.

A third key intellectual difference between the approaches stems from differences in how they unify phenomena. Unlike the elementary approach, both symmetry-based approaches unify the crystal field theory problem into symmetry-based families.[15] The symmetry-based derivations apply not only to a given coordination complex but also to

[15] The elementary approach cannot unify the energy-level structure because it treats each coordination complex on a case-by-case basis. This piecemeal approach results from its central EDR, namely that knowledge of the eigenvalues is sufficient for knowledge of the splitting and degeneracy. To calculate these eigenvalues, it—like the other two approaches—appeals to the particular strength of the interaction, a feature specific to each coordination complex.

coordination complexes in the same geometric family. Specifically, they provide the following kind of epistemic dependence relation: knowledge of the energy-level structure for one coordination complex in this family suffices for knowledge of the energy structure for other complexes in this family. Unifying coordination complexes into symmetry-based families enables us to understand each one as an instance of a larger class with the same behavior. Moreover, due to the differences in their respective EDRs, the group-theoretic approach unifies more than the non-group-theoretic approach. In the latter, a given symmetry-based potential applies only to coordination complexes that have the same geometric arrangement of ligands around the central metal ion. For nickel(II) hexahydrate, this is a particular instantiation of octahedral symmetry, with each ligand at the vertex of an octahedron. In contrast, the relevant group-theoretic argument applies to any coordination complex with octahedral symmetry, independently of how it is instantiated. For instance, the derivation sketched in Section 4.2.1 applies just as well to a coordination complex with eight ligands at the vertices of a cube, rather than an octahedron.[16] The group-theoretic approach unifies more because it tells us that knowledge of the abstract symmetry group suffices for knowledge of the energy-level structure.

Despite my focus on non-pragmatic differences in understanding, some EDRs are pragmatically beneficial as well. For agents interested in knowing only the splitting and degeneracy, modularization provides a beneficial way of obtaining this knowledge without needing to determine further properties of the eigenvalues. Similarly, for agents who wish to determine the energy-level structure of a class of coordination complexes, the additional unification provided by group theory hastens this task. In this way, different EDRs can lead to pragmatic benefits, based on agents' goals. But, unlike many recent pragmatic accounts of scientific understanding, conceptualism keeps these two aspects of theory reformulation separate, i.e. the pragmatic benefits vs. the objective and non-pragmatic differences in understanding. The former depend on the preferences and goals of agents, whereas the latter are formal properties of a theory's formulation.[17]

[16]Since cubes are dual to octahedra, they have the same symmetry group.

[17]Against my approach, Potochnik (2015a, p. 1172) argues that it is a mistake to seek a clean-divide between the pragmatic and the non-pragmatic in the context of explanation and understanding. Defending the utility and coherence of this basic distinction—which explanationists necessarily grant—lies outside the scope of this chapter.

4.3 Is Conceptualism Redundant?

Section 4.2.2 developed a counterexample to explanationism. This counterexample targets explanationists who view each approach to crystal field theory as explanatory—an interpretation sanctioned by many accounts of explanation. To rebut my argument, these explanationists might argue that other non-pragmatic features of scientific explanation—such as differences in explanatory depth—account for all intellectual differences between the approaches. This would render conceptualism's account of scientific understanding redundant, relative to existing accounts of explanatory differences.

In response, I will argue that conceptualism is not redundant when compared with leading accounts of explanation that treat the approaches to crystal field theory as explanatory. Because I view them as the most promising theories in their respective traditions, I will focus on Skow's (2016) account of reasons-why and Woodward and Hitchcock's manipulationist account.[18] I will argue that both accounts fail to accommodate modularization and unification. Additionally, even Kitcher's (1989) unificationist account of explanation does not accommodate the relevant kind of unification illustrated by crystal field theory.

4.3.1 Skow's account of reasons-why

Skow's account of reasons-why continues the causal explanation tradition of Railton and Lewis, while treating grounding as an additional explanatory feature. For a given concrete event Q, Skow's theory characterizes a hierarchy of reasons why Q occurred. At bottom, there are the 'first-level reasons why Q.' These are always either causes or grounds. For each reason-why, there might be further reasons why that reason is a reason, and so on (2016, p. 124). Nevertheless, when it comes to answering the initial, bottom-level why-question regarding Q, answering these higher-level why-questions is optional on Skow's account. Hence, explanatory arguments that agree on the first-level reasons-why provide the same explanation. Since the three approaches to crystal field theory agree on the un-

[18]Potochnik's causal pattern account of explanation provides an interesting approach that amalgamates and develops aspects of Woodward's, Achinstein's, Strevens', and van Fraassen's accounts of explanation (2015a, 2017, pp. 127, 134). However, since it relies on pragmatic features of agents (2015a, pp. 1172–5, 2017, p. 127), it is not amenable to a defender of explanationism. Bokulich's (2011) rich account of model explanations modifies Woodward's account to accommodate idealization, a complication that I suppress here.

derlying ontic reasons for the crystal field theory phenomena, Skow's account would treat each of them as explanatory.[19] Thus, his account falls within the scope of Section 4.2.2's argument against explanationism.

Although Skow's account has its attractions, it is too circumscribed to capture many of the intellectually significant differences brought out by symmetry arguments. The chief utility of Skow's theory derives from characterizing a hierarchy of reasons-why. This enables us to see how multiple arguments can agree on a set of lower-level reasons-why while differing in the higher-level reasons that they articulate. However, Skow's theory focuses exclusively on answers to why-questions, and modularization is not an answer to a why-question. Instead, modularization is a feature of how some explanatory arguments are organized, based on the epistemic dependence relations they deploy. As a property of EDRs, modularization characterizes what it *suffices to know* to answer a certain why-question. Characterizing what-it-takes-to-know something is different in kind from answering a why-question about a physical phenomenon.

For a different reason, Skow's account cannot accommodate the symmetry-based unification discussed in Section 4.2.3. Skow restricts his account to reasons why to 'concrete events,' excluding law-like generalizations such as Galileo's law of freefall (2016, pp. 27–8, 37). This restriction makes it difficult to explain generalizations, such as the claim that all coordination complexes with octahedral symmetry display a particular energy-level structure. To accommodate such generalizations, Skow's account must treat them as a conjunction of concrete events, explained by a concatenation of reasons-why for each system (2016, p. 134). This sort of aggregative explanation leaves open that these different systems only coincidentally display the same pattern of behavior. Aware of this worry, Skow claims that "to show that it is no coincidence that all the facts in some collection obtain it is enough to find a *common reason* why they all obtain" (2016, p. 134). In this case, the common reason would presumably be that all these different coordination complexes have octahedral symmetry. Nevertheless, as Lange (2010, pp. 307, 319–22, 2014, pp. 508–9) notes, there is an important difference between conjoining explanations (even those sharing a common explainer) vs. providing a single, unified explanation. Skow's account cannot afford any special significance to the single, unified derivation that group

[19] Technically, Skow recommends abandoning the explanation-idiom in favor of answers to why-questions that describe the reasons why an event occurs (2016, pp. 7–10). Nonetheless, 'explanation' remains a convenient catchall for the particular kinds of reasons-why and why-questions pertinent to science.

representation theory provides. For indeed, the non-group theoretic symmetry argument provides the same common reason—octahedral symmetry—but without providing a single unified derivation that covers all instantiations of this symmetry group. Hence, Skow's account would have to view these two approaches as being on a par with respect to unification, although they are not.

4.3.2 Woodward and Hitchcock's manipulationism

Manipulationism also fails to accommodate modularization and unification. As described in Section 4.2.2, manipulationism focuses on answering what-if-things-had-been-different questions, using possible interventions on the system of interest. However, just as modularization is not a reason-why, it is also not subsumed under answers to what-if questions—again because modularization is an organizational feature of explanatory arguments. Thus, it *prima facie* lies outside the scope of manipulationism. In particular, manipulationism neglects a crucial instance of modularization, namely the EDR that we do not need to know the charges and distances in order to determine the splitting and degeneracy. The group-theoretic approach provides us with this EDR by demonstrating that knowledge of symmetry suffices for knowledge of the energy structure. However, manipulationism cannot accommodate this EDR using interventions and what-if-things-had-been-different questions. The problem stems from Woodward and Hitchcock's stricture that explanations depend solely on "invariance under some range of changes *in the variables figuring in the [explanatory generalization] itself*" (2003a, p. 7). Changes to other variables left out of the explanatory generalization—such as background conditions—do not figure in the explanation. Crucially, the group-theoretic approach gives us the above EDR by setting aside explicit dependence on the charges and distances, treating them as background conditions. Hence, when it comes to explaining the splitting and degeneracy, the group-theoretic approach is silent on interventions that affect these variables.[20]

Nevertheless, an explanationist might use manipulationism's account of explanatory

[20]Woodward does discuss a completely different notion of 'modularity' in the context of representing causal structure by systems of equations (2003, pp. 48, 327–9). This notion of modularity requires that each equation represents a distinct causal mechanism, so that we can intervene on one equation without affecting others. The crystal field theory equations are not modular in this sense because they are not causally independent of each other.

depth to underwrite these intellectual differences. According to Woodward and Hitchcock (2003a, 2003b), one generalization is—*ceteris paribus*—deeper than another if the former incorporates an explanans that the latter treats as a background condition. Explicitly incorporating background conditions shows how the phenomenon depends on additional explanans, thereby providing a deeper explanation. For instance, the group-theoretic character decomposition formula (Equation A.5) explicitly incorporates symmetry as an explanans variable. This makes the group-theoretic approach deeper than the non-group-theoretic approach because the latter fixes symmetry as a background condition. While this might be a welcome result, Woodward and Hitchcock's account of depth also has counterintuitive consequences. In particular, it seems to classify the elementary approach as deeper than the group-theoretic approach: the elementary approach allows for interventions on not only the symmetry but also the charges and ligand distances, whereas the group-theoretic approach treats these latter features as background conditions. Yet, the group-theoretic EDRs tell us something important about crystal field theory that the elementary approach neglects. Independently of whether either approach is deeper, we need an account of this intellectual difference. This is what conceptualism provides.

Moreover, Woodward and Hitchcock explicitly disavow that unification matters for causal explanation. Recall that symmetry arguments unify by focusing on families of systems that share a key property, such as the family of coordination complexes with octahedral symmetry. Woodward and Hitchcock argue that this kind of generalization is not crucial for accounts of explanation at all. Instead, they focus exclusively on a second kind of generalization, based on varying the properties of a particular system. They characterize this as "generality with respect to *other possible properties of the very object or system that is the focus of explanation*" (2003b, p. 182).[21] Having specified a particular system of interest, one considers varying features that are properties of that system only. They thus deny any need to interpret unification *qua* explanation. As Woodward argues, many kinds of unification involve classificatory schemes or general mathematical formalisms that are not intrinsically connected with causal explanations (2003, pp. 362–4). This illustrates how accounts of explanation can neglect intellectually significant features such as unification. Not everything that matters for understanding necessarily has to do

[21]Likewise, Woodward claims that "the explanatory depth of a generalization is connected to its range of invariance rather than its scope; hence, the unificationist approach focuses on the wrong sort of generality in explanations" (2003, p. 366).

with explanation.[22]

Still, an explanationist might wonder whether Kitcher's (1989) account of unification already accommodates the intellectual differences wrought by unification. Perhaps surprisingly, it does not. According to Kitcher, an argument pattern counts as explanatory provided that it best unifies the phenomena. This is quantified in terms of deriving the largest number of phenomena relative to the smallest number of assumptions. Only these unificatory argument patterns earn a place in the 'explanatory store' of arguments that are genuinely explanatory. But, problematically for Kitcher, the explanatory store is deductively closed, and this prevents it from distinguishing the three approaches to crystal field theory on unificatory grounds. Recall that the group-theoretic approach relies on the same perturbation-theoretic argument schema as the other two approaches. Hence, including the group-theoretic approach within Kitcher's explanatory store *ipso facto* includes the other two approaches. Thus, even though only the group-theoretic approach maximally unifies crystal field theory, Kitcher's account does not distinguish it from the other approaches. This shows that Kitcher's notion of unification is actually too weak to adequately characterize the relevant EDRs that I have identified.[23]

4.4 Explanatory Exclusion

Short of denying that there are intellectual differences between compatible formulations, seemingly only one strategy for defending explanationism remains: an explanationist could deny that the three formulations of crystal field theory are each explanatory. If only the group-theoretic approach provides a genuine explanation, then its intellectual differences would arise from explanatory differences after all. Mounting this strategy requires adopting an *exclusionary account* of explanation. Compared to the accounts considered in Section 4.3, exclusionary accounts posit more restrictive criteria for explanatory relevance. By making the criteria for explanation more demanding, exclusionary accounts generate an explanatory difference between compatible formulations. These putative explanatory differences can ground corresponding differences in understanding, thereby

[22] See Gijsbers (2013) for a similar conclusion. Gijsbers (2007) provides a detailed argument for why unification is not inherently connected with explanation. Similarly, Morrison (2000) argues through numerous case studies that unification is often either in tension with or has nothing to do with explanation.

[23] For additional criticisms of Kitcher's account, see Barnes (1992) and Woodward (2003, pp. 366–73).

precluding counterexamples to explanationism.

To illustrate this strategy, I will consider two exclusionary accounts: Strevens' (2008) kairetic account and Lange's (2017) account of distinctively mathematical explanations. Both accounts focus on abstraction of allegedly irrelevant causal details, thereby excluding many causal influences from counting as explanatorily relevant. We will see that only the group-theoretic approach satisfies their restrictions, making it the only genuine explanation for crystal field theory. However, this exclusionary strategy runs afoul of the third desideratum from Section 2.3: ideally, we should accommodate the apparent differences in understanding without appealing to epistemically inaccessible ontic features. Positing more restrictive explanatory relevance relations generates a skeptical problem because we cannot easily know whether these additional relevance relations exist or are satisfied.[24] My goal here is not to reject these accounts of explanation *per se*, but rather to point out that no one should adopt them merely for the sake of upholding explanationism. Conceptualism provides a more epistemically secure and parsimonious account of the relevant differences in understanding.

4.4.1 Strevens' kairetic account

Strevens' account focuses on identifying *causal difference-makers*, the only causal influences that are explanatorily relevant. According to his 'kairetic condition,' a causal influence counts as a difference-maker provided it remains in at least one maximally abstract model explaining the phenomenon. To apply this test, we begin with a model that causally entails the explanandum.[25] We then make this causal model as abstract as possible, replacing specific descriptions of causal influences with increasingly abstract characterizations, i.e. less exact or specific claims (Strevens 2008, p. 97). Causal influences that survive this abstraction procedure qualify as difference-makers. On Strevens' account, only these maximally abstract causal models genuinely explain.[26]

Applying the kairetic account to the three formulations of crystal field theory, we see this abstraction procedure in action. Regarding the splitting and degeneracy, the non-

[24]Woodward defends a similar epistemic accessibility criterion for explanation (2003, pp. 23, 179–81, 308).

[25]Causal entailment goes beyond logical entailment by representing an actual causal process that produces the explanandum (Strevens 2008, pp. 71–2, 93). Potochnik weakens this entailment relation to better accommodate idealizations (2017, pp. 155–6).

[26]Note that a single application of this kairetic procedure only identifies all of the difference-makers that appear in a given causal model, rather than all of the difference-makers for a given event.

group-theoretic approach abstracts from the particular charges and field strength used by the elementary approach.[27] In their stead, it offers the symmetry-based potential as a putative causal difference-maker. The group-theoretic approach abstracts further, eliminating the particular way that symmetry is instantiated. Assuming that we can always re-express knowledge of symmetries using group theory, the symmetry-based potential would be explanatorily irrelevant. What remains are the symmetries themselves—of the initial metal ion and the resulting coordination complex—as the putative causal difference-makers for the splitting and degeneracy. Since only the group-theoretic approach successfully represents these difference-makers, it is the only approach we have considered that would provide a genuine explanation. It shows that symmetry is a difference-maker, but not the particular way that symmetry is instantiated. In this way, Strevens' kairetic account could ground the intellectual differences between the two symmetry-based approaches in corresponding ontic explanatory differences, thereby preserving explanationism.

However, the kairetic abstraction procedure for identifying causal difference-makers faces a skeptical challenge. In general, it is impossible to conclusively prove that any causal factor is or is not a difference-maker. On the one hand, to show that a factor is a difference-maker, we must show that it survives under maximal abstraction within a causal model. Yet, how can we know that no further abstractions in our model are possible? A scientist taught the non-group-theoretic approach—and with no knowledge of group theory—might reasonably think that this is a maximally abstract causal model. On the other hand, to show that a causal factor is not a difference-maker, we must show that the kairetic procedure eliminates or abstracts away that causal factor from any and all causal models for the given explanandum (Strevens 2008, pp. 69–70, 87). For instance, to show that the specific eigenvalues are genuinely irrelevant for explaining the splitting and degeneracy, it is not enough to see how group theory eliminates them to provide a more abstract causal model. Instead, we would have to show that any causal model for the splitting and degeneracy lets us abstract away the specific eigenvalues.

Conceptualism avoids these skeptical worries by analyzing the relevant intellectual differences in terms of epistemic dependence relations, rather than putative causal difference-makers. It shows that we do not have to consider other possible but currently

[27] We cannot, however, conclude from this single application of the kairetic procedure that the charges and field strength are explanatorily irrelevant for the splitting and degeneracy. To do that, we would have to show that we can abstract them away from any model that causally entails the splitting and degeneracy.

unconceived causal models to account for the change in understanding provided by group theory. Although considering such models would no doubt be illuminating—since it would amount to considering further reformulations—conceptualism lets us analyse the intellectual differences between the symmetry-based approaches by considering those two approaches alone. The hard task of identifying causal difference-makers may reasonably be hostage to the existence of even more abstract models, but the task of identifying intellectually significant differences surely is not.

4.4.2 Lange's distinctively mathematical explanations

Faced with the limitations of causal accounts of explanation, an explanationist might try to locate ontic differences between reformulations within the realm of non-causal explanations. Perhaps what is needed to save explanationism is an account of how mathematical facts explain physical phenomena. Lange's theory of distinctively mathematical explanations provides one such account. According to Lange, causal structure alone sometimes cannot account for the inevitability of certain physical phenomena (2017, pp. 5–6). In these cases, we require a 'distinctively mathematical explanation,' wherein a mathematical fact 'constrains' the causal structure of reality. Such constraints possess a higher degree of necessity than the laws or contingent facts that they constrain (Lange 2013, 2017, p. 10). Recognizing the relevant constraints shows not only why the explanandum occurred, but also why the explanandum was inevitable—in a modal sense stronger than nomic inevitability. In such cases, arguments that appeal to causal structure alone are merely derivations of the relevant phenomenon, rather than explanations.

When it comes to explaining splitting and degeneracy, Lange's account classifies only the group-theoretic approach as explanatory. On this interpretation, the character decomposition formula (Equation A.5) constrains the possible form of all resulting energy levels, given initial and final symmetry groups. In other words, it constrains all possible causal relations governing the energy-level structure of coordination complexes. For instance, if the force law governing coordination complexes weren't Coulomb's law—e.g. if it were an inverse cubic force law instead—the splitting and degeneracy would remain the same. Although naturally necessary, such force laws would be interpreted as less necessary than the mathematically necessary facts governing the representations of symmetry groups. Hence, the group-theoretic approach shows that the splitting and degeneracy are more necessary than both the relevant force laws and the resulting eigenvalues. Since the other two approaches cannot explain this difference in modality, Lange would argue that they are non-explanatory: the causal mechanisms they reference are explanatorily irrelevant.

However, just like Strevens' account, Lange's account faces a skeptical challenge. His account successfully distinguishes the group-theoretic approach from the other approaches only if the world possesses this rich modal structure of mathematical facts constraining physical facts. Commitment to this kind of graded modality is decidedly controversial, in part because it is epistemically inaccessible. We do not have empirical access to this hierarchy of modal facts. Nevertheless, Lange's account applies only if the world possesses this structure. If the world turned out to lack sufficient modal structure, then distinctively mathematical explanations would devolve into ordinary causal explanations.

A further problem arises from this skeptical worry. Worlds that lack Lange's requisite modal structure are empirically indistinguishable from worlds that possess it. Hence, the intellectual differences described in Section 4.2 would be equally apparent in either kind of world. Regardless of whether or not the world possesses this modal structure, the three formulations would still provide different understandings of crystal field theory. Since these differences do not depend on corresponding facts about modality, we should be able to accommodate them without further metaphysical theorizing. For instance, to appreciate the central insight that modularization provides, we don't need there to be graded modality in the world. Specifically, we don't need group-theoretic facts to constrain the causal structure of the world. By supplying epistemic dependence information, the group-theoretic approach makes a distinctive contribution to our understanding of crystal field theory, independently of further ontic commitments. Hence, conceptualism provides a superior strategy for accommodating these intellectual differences.

Plausibly, any exclusionary account of explanation will face similar problems. For the basic idea behind this explanationist strategy is to posit additional ontic features that might ground the intellectual differences between compatible formulations. Since compatible formulations superficially posit the same states of affairs, any such additional ontic features will be epistemically less accessible. In contrast, conceptualism provides an account of these intellectual differences using features that we have easy epistemic access to—namely, epistemic dependence relations. Determining epistemic dependence relations is simply a matter of analyzing the epistemic structure of a theory formulation. Since these differences in EDRs persist independently of whether they are grounded in further ontic differences, our account of understanding should likewise be independent of these further differences. Whereas conceptualism satisfies this desideratum, the exclusionary strategy canvassed here seemingly does not.

4.5 Expressivism about Comparative Understanding

So far, I have offered an account of *differences in understanding*. These differences concern how one formulation provides a different understanding of a phenomenon than another. There is a further interesting question: under what conditions does one reformulation provide a *better understanding* than another? On what grounds—if any—is one compatible formulation "intellectually better" than another? In asking such questions, we implicitly ascribe a kind of intellectual value to our problem-solving procedures. We entertain the possibility that one of these procedures has greater intellectual value than another, thereby providing better understanding. What is the nature of this intellectual value that underpins claims about comparative understanding?

In keeping with the minimal ontological commitments of Chapter 2, I am reluctant to posit metaphysically robust facts about intellectual value. The natural or non-natural world might come equipped with such facts, but we should try to avoid positing them if we can. Otherwise, we face problems of epistemic access similar to those that plague fundamentalism. Fortunately, there is a general philosophical strategy for simultaneously vindicating a discourse while avoiding commitment to mysterious entities or facts that such a discourse implicitly references. In its contemporary guise, this strategy is known as 'expressivism' or 'quasi-realism.' It is most well-known in ethics, where expressivism allows us to recover and vindicate ordinary moral discourse without positing metaphysically robust moral facts or properties.[28] Here, I will propose an expressivist account of comparative understanding.

Traditionally, expressivism begins with a distinction between descriptive vs. non-descriptive claims. Descriptive claims are what Field (2009) calls *straightforwardly factual.* They represent states of affairs, mirroring reality. In contrast, non-descriptive claims

[28] For developments of quasi-realism in the ethical domain, see Gibbard (1990) and Blackburn (1993).

are not straightforwardly factual; they involve an evaluative or normative dimension.[29] Expressivism focuses in particular on those non-descriptive claims that express action-directed states of mind (Chrisman 2007, p. 236). Various expressivist accounts differ on the mental states or attitudes involved, but common ones include universal preferences, states of norm-acceptance, states of planning, or a non-cognitive attitude of being for.[30] Because I find them complementary and illuminating, I present my account using both Schroeder's (2008) attitude of being-for and Gibbard's (1990) attitude of norm-acceptance.

4.5.1 A first go at expressivism about understanding

Judgments of comparative understanding are a type of evaluative judgment. They are judgments that one formulation provides *better understanding* than another. More precisely, they take the form "X provides a better understanding than Y of some phenomenon P." The evaluative part comprises a "better-than" relation. Hence, if we can characterize what it means to say that "X is better than Y," we will have gone quite a ways toward characterizing comparative understanding.

This is one place where expressivism gains a clear foothold. Expressivism provides a general strategy for understanding better-than relations. Generically, to judge that something X is better than some other thing Y is to express an attitude of *being for preferring X to Y* (Schroeder 2008, p. 58). Notice how this expressivist account differs from a descriptivist (or representationalist) account of 'better-than.' Descriptivism would posit some state of affairs that 'better-than' relations track, such as comparative amounts of psychological well-being. Expressivism can remain neutral on whether 'better-than' relations ever track such states of affairs. The point is a modest one: we can at least understand 'better-than' claims through their expressive role. They at least express an attitude of being for preferring or favoring one thing over another.

Naïvely, expressivism about better-than relations leads to the following account of comparative understanding:

> **(Naïve) Comparative Understanding**: to judge that an argument (or problem-solving plan) X provides better understanding than an argument Y is to express an attitude of *being for preferring X to Y.*

[29]See Kraut (1990, p. 159) for a detailed characterization of this 'bifurcation' between descriptive and non-descriptive claims and Price (2013) for criticism.

[30]See Gibbard (2003, p. 181). Schroeder's (2008) account in terms of the attitude of 'being for' resolves some technical problems that afflict Gibbard's solution to the Frege–Geach embedding problem.

On this simple account, when I judge that the group theoretic approach to crystal field theory provides a better understanding than the elementary approach, I express an attitude of being for preferring group theory. Here, the object of understanding can be a specific phenomenon, proposition, or body of knowledge.

Of course, there are many different kinds of preference. Some preferences are intended as expressions of taste, whereas others are intended to express more universalizable attitudes. When I say that chocolate ice cream is better than vanilla, I don't intend my expression of preference to be anything more than a matter of taste. I am communicating that, as a matter of my contingent physiology, I find chocolate to be more enjoyable than vanilla. I am not encouraging others to agree with me; I do not care about their preferences on this matter.

In contrast, when I say that helping is better than hurting, I take this expression of preference to be more than a matter of taste. I express an attitude of preferring that *everyone* think helping is preferable to hurting. I express a preference that I think everyone should have. Clearly then, to make the account work, we need to say more about the kinds of preferences involved.

4.5.2 Intellectual preference

The naïve account of comparative understanding runs immediately into a serious problem: many preferences are idiosyncratic, subjective, or practical. As we have seen in the context of trivial notational variants, I might have a strong preference for working with one notational convention rather than another. I might strongly prefer working in a right-handed rather than a left-handed coordinate system. Nevertheless, as I have argued in Chapter 2, either convention results in the same understanding of the problem. It is in this sense that they are trivial notational variants: they evince no non-practical, epistemic differences. *A fortiori*, neither can provide a better understanding than the other, at least not in the sense of 'understanding' that interests me here.

To avoid this problem, it suffices to restrict the relevant notion of preference. Rather than focus on all of an agent's preferences toward formulations, I will focus on their *intellectual preferences*. These comprise their preferences toward intellectual differences or features that give rise to such differences (recall that 'intellectual differences' constitute a non-practical subset of epistemic differences). These include preferences toward modularization, unification, uniformity of treatment, and manifestness. Since intellectual differences arise from differences in epistemic dependence relations, it follows that intellectual preferences also include preferences toward specific EDRs. Given any two epistemic dependence relations that can be used in solving a problem, I might have an intellectual preference for one of them over the other.

In contrast, intellectual preferences do not distinguish between trivial notational variants. Since a right-handed coordinate convention does not yield any intellectual differences from a left-handed system, my preference for right-handed systems is not an intellectual preference. It is what we might call a 'practical preference'—stemming from my familiarity or relative facility with right-handed systems over left-handed ones. There are typically many practical reasons to prefer one formulation over another, based on familiarity with a formalism, preferences for the kinds of computations one likes to perform, mathematical background knowledge, etc. These practical reasons can contribute to overall preference, but not intellectual preference.

By restricting attention to intellectual preferences, we arrive at the following improved account of comparative understanding:

> **Comparative Understanding**: to judge that an argument X provides better understanding than Y is to express an attitude of *being for intellectually-preferring X*.

On this account, when I judge that the group theoretic approach provides a better understanding of crystal field theory than the elementary approach, I express an attitude of intellectually favoring the group theoretic approach. Specifically, I might be expressing an intellectual preference for modularization and unification. These are two non-practical, epistemic features that the group theoretic approach supports that the elementary approach does not. Often, to judge that formulation X provides better understanding than formulation Y will amount to judging that X provides, on balance, a better set of EDRs than Y. Moreover, to judge that these EDRs are better is to express an attitude of being for intellectually preferring them.

Of course, in various contexts, my overall preference might be for the elementary approach. Someone who is interested in solving only a few crystal field theory problems might prefer, for the sake of convenience, to avoid learning any group theory at all. If so, they might overall prefer to use the elementary approach, which avoids the hard work of tabulating character tables for the relevant symmetry groups. However, this kind of convenience-based preference is not relevant for an assessment of comparative intellectual virtues. Focusing on intellectual preferences screens off matters of convenience from matters of intellectual importance.

4.5.3 Norm-expressivism about better understanding

So far, I have relied on Schroeder's formulation of expressivism, using a non-cognitive attitude of *being for*. Schroeder's formulation provides a relatively streamlined characterization of the attitudes that judgments of comparative understanding express. Nonetheless, focusing on norms provides a richer account of these judgments. We can unpack the attitude of *being for* in terms of accepting a set of norms, i.e. through an attitude that Gibbard (1990) calls 'norm-acceptance.' On this construal, 'being for X' amounts to being in favor of a set of norms that *recommend X*.

I intend this notion of 'recommendation' to be a weak form of 'ought,' in the sense of 'you ought to take out your trash,' rather than the stronger sense of 'you must take out your trash.' Recommendation, in this sense, is stronger than permissible but weaker than required. Although Gibbard (1990) talks in terms of actions being required, permissible, or forbidden, I see no reason why we can't adjoin a further normative notion of recommendation. Someone who fails to follow a recommendation engages in an action that, although permissible, is suberogatory.

It is very natural—at least in a philosophical context—to talk about norms governing better understanding. These norms govern what counts as an intellectual improvement, i.e. improvements to our understanding in a non-practical, epistemic sense. Hence, it is illuminating to reformulate my account of comparative understanding using Gibbard's norm-expressivism. I take Gibbard's framework to be compatible with Schroeder's, and I do not intend to treat either as more fundamental. Rather, both frameworks elucidate different facets of comparative understanding (and evaluative judgments more generally).

Gibbard's norm-expressivism

Gibbard's framework extends possible worlds semantics by adjoining a set of norms to each possible world. The worlds are specified entirely through descriptive claims, i.e. claims that purport to represent states of affairs or mirror reality. Adjoining a set of norms yields descriptive–normative worlds, given by a pair $\langle w, n \rangle$.[31] The possible world w specifies states of affairs. The additional component n specifies a normative system. Gibbard characterizes this as follows:

> A system of norms, recall, is the end result of the ways the various general normative principles a person accepts combine, weigh against each other, and override one another. If it is complete, then for every conceivable fully described occasion governed by norms, the system classifies each alternative as required, optional, or forbidden....Together, w and n entail a normative judgment for every occasion. (Gibbard 1990, p. 95)

To illustrate Gibbard's framework, consider judgments about rationality. When a norm-expressivist about rationality says that X is *rational*, they neither ascribe a property to X nor assert a truth-condition for X (at least not directly). Instead, saying that "X is rational" is equivalent to expressing acceptance of a system of norms that, on balance, permit X (Gibbard 1990, p. 84). In short, to think something rational is to accept a set of norms that permits it. For any given belief, judgment, or action, a complete system of norms renders it either required, permissible, or forbidden. Below, I will further divide permissible states into those that are recommended and those that are not.

Applying Gibbard's Framework

To apply Gibbard's norm-expressivism to understanding, I must specify which components fall within the descriptive part, and which fall within the normative part. Intellectual differences belong to the descriptive part: it is a matter of straightforward fact whether or not two formulations display intellectual differences. In saying that the group theoretic approach modularizes the crystal field theory problem, while the elementary approach does not, I say nothing normative. I merely describe non-practical, epistemic features of these two formulations.

The normative part comprises norms governing intellectual preferences. These norms settle which intellectual preferences are recommended, required, forbidden, or permissible-but-not-recommended (i.e. suberogatory). For instance, one such norm might say that, *ceteris paribus*, one ought to intellectually prefer unification. According

[31]Gibbard himself refers to these as 'fact–norm' pairs, but due to his later embrace of minimalism about truth, this terminology is liable to cause confusion. Within the minimalist framework, normative claims can also count as being factual (although not straightforwardly factual).

to this norm (considered in isolation), a formulation that unifies provides better understanding than one that does not, assuming that they are otherwise on a par intellectually (i.e. with respect to other intellectually significant features). When I say that the unifying formulation provides better understanding, I express acceptance of a set of norms that recommends this intellectual preference.

(*An aside*: technically, matters are not so simple for *all* intellectually significant properties. Some of these properties themselves have a normative component and hence are not wholly descriptive, i.e. straightforwardly factual. As I have mentioned, making a property manifest is an intellectually significant feature that some formulations possess while others lack. Yet, as I argue in Chapter 6, what it means to make a property manifest is partly normative. Whether or not a property is manifest depends on what we *epistemically ought* to infer from a given expression or argument. Nonetheless, the kind of normativity involved in these epistemic ought claims is different from the normativity involved in 'better-than' relations. Epistemic-oughts have nothing to do with intellectual preferences. If one were a descriptivist about epistemic-ought claims, then the simple story in the preceding two paragraphs would hold.)

Within this framework, disagreements about comparative understanding are either descriptive or normative. They amount to disagreements about either i) which formulations display a given intellectually significant feature or ii) the right norms governing intellectual preferences. Imagine then two chemists, Greg and Eleanor, who disagree about whether the group theoretic approach provides better understanding than the elementary approach. Stipulate that they both agree on the intellectual differences between these approaches. In particular, they both recognize that only the group theoretic approach modularizes and unifies the crystal field theory problem. According to Greg, the group theoretic approach provides better understanding, whereas Eleanor denies this (perhaps thinking that both provide an equally good understanding).

In making his comparative judgment, Greg expresses acceptance of a set of norms that recommends intellectually preferring the group theoretic approach. In denying that group theory provides better understanding, Eleanor expresses acceptance of norms that forbid Greg's intellectual preference (or at least recommend a different preference). In thinking that at most one of Greg or Eleanor can be correct, I myself express a further normative judgment. I express acceptance of a set of norms that forbids both being right.

In other words, I express my rejection of relativism about norms governing intellectual preference. In this way, expressivism does not entail relativism: relativists about comparative understanding endorse a particular set of norms governing intellectual preferences. To avoid relativism, it suffices to reject such norms.

It is of course possible that in some cases, the right thing to say is that neither formulation provides a better understanding than the other. The two formulations might simply display different intellectual virtues that balance each other out. For instance, perhaps the correct norms on intellectual preference do not legislate in favor of one epistemic feature over the other. In such cases, the most we might be able to say is that it is ideal to be able to understand the phenomena using both approaches, rather than to intellectually privilege one over the other. Section 4.6.3 further explores the possibility of there being no privileged weighting of various intellectually significant properties.

4.5.4 Expressivism is not subjectivism

Of course, we should expect that different scientists and scientific communities in various contexts and historical periods will disagree about which epistemic dependence relations should be preferred. When it comes to making sense of these disagreements, expressivism improves upon agent- or community-based subjectivism. According to subjectivism, judgments of comparative understanding are *synonymous with* judgments of preference. If a scientist Greg says that "group theory provides a better understanding of crystal field theory than the elementary approach," a subjectivist understands this as follows: Greg prefers group theory to the elementary approach.

Subjectivism faces serious problems for making sense of disagreements. If a different scientist Eleanor asserts that the elementary approach provides better understanding, a subjectivist construes this as synonymous with the following assertion: "Eleanor prefers the elementary approach to the group theoretic approach in the context of crystal field theory." Problematically, both of these assertions about Greg and Eleanor's preferences can be true simultaneously. Hence, subjectivism seems unable to capture a sense in which Greg and Eleanor disagree. It is unable to capture the ordinary intuition that Eleanor not only disagrees with Greg's attitudes or preferences, but also thinks that the content of his claim about comparative understanding is false.[32] In contrast, on an expressivist

[32]For discussion of this problem for speaker subjectivism, see Schroeder (2008, 16ff.). Beddor (2019)

account of comparative understanding, Greg and Eleanor are in genuine disagreement. They disagree about which epistemic dependence relations we ought to prefer.

There are good historical reasons to resist subjectivist approaches to comparative judgments of understanding. The mere consensus of the scientific community is insufficient evidence for treating one formulation as intellectually better *tout court* than another. Forming consensus about the betterness of reformulations generally takes time and depends largely on training. It is subject to numerous historical contingencies, including the idiosyncratic predilections of individual scientists and promoters of favored reformulations. Overall preferences are also heavily influenced by which approach is ultimately more convenient for a class of problems. As discussed briefly in Section 2.4.1, even deciding which of two approaches is more convenient can depend on personal preferences. For instance, although group theoretic methods ultimately lead to vast computational simplifications, they have a steep entry cost compared to elementary approaches. Learning the requisite group representation theory is not easy compared to applying elementary brute-force methods. It is largely for this reason (along with the pedagogical absence of group theory in early 1900s physics education) that group theoretic approaches to atomic spectroscopy took more than two decades to become widespread.[33]

Unlike subjectivism, expressivism avoids making better-understanding a matter of historical contingency and subjective preference. Within an expressivist framework, we can criticize scientists from various periods for having poor intellectual preferences about various kinds of epistemic dependence relations. Additionally, we can criticize them for conflating their overall preference (taking into account factors of convenience) with properly intellectual preference. Chemists and physicists have since come to appreciate the intellectual advantages of group representation theory. These advantages were there all the while to be admired, independently of whether chemists and physicists ever did admire them. Thus, expressivism allows that many disputes about better-understanding are substantive intellectual debates rather than matters of taste. It is plausible that epistemic dependence relations that modularize and unify are ones that scientists ought to strive for, *ceteris paribus*. If so, contemporary chemists and physicists would have a principled reason to criticize the preferences of many of their forebears in the 1930's and 1940's.

discusses how expressivism improves upon subjectivism and contextualism.

[33]For historical accounts, see Scholz (2006) and Bonolis (2004)

4.6 Norms on Better Understanding

By itself, expressivism does not provide a first-order account of the norms governing comparative understanding. In the first instance, expressivism is a meta-theoretical framework: it clarifies what we are doing when we make certain kinds of non-descriptive claims. Providing a first-order account requires endorsing and defending a particular set of norms, including which EDRs and kinds of intellectual properties to prefer. In the ethical domain, expressivism is compatible with a wide variety of first-order moral theories. An expressivist could endorse utilitarianism or deontology or Aristotelian virtue ethics. Likewise, my expressivist account is compatible with a variety of different normative systems governing comparative understanding. Nonetheless, I recognize that many philosophers are most interested in determining these norms, since they settle what actually contributes to greater understanding. This section describes different first-order norms on comparative understanding.

I divide these norms into three families. The first stems from the relationship between explanation and understanding. I imagine that these norms will be the least controversial, since they are entailed by a widely-accepted account of explanatory understanding (i.e. understanding-why). They are for that reason relatively uninteresting. The second family focuses on improvements coming from increasing the number of epistemic dependence relations or intellectually significant properties. The third family is the most interesting: it concerns norms governing which non-explanatory intellectually significant properties to prefer, including modularization, unification, and manifestness. I imagine that reasonable people will probably be most likely to disagree about these norms. I myself am inclined to endorse all of the norms that I consider. I provide brief arguments for some of these norms, although part of the defense for the third family of norms must await Chapter 5.

4.6.1 Norms from explanation

According to *the received view of understanding*, understanding why a phenomenon P occurred consists in grasping an explanation of that phenomenon. This account entails an uncontroversial claim about comparative understanding: given two arguments, if only one of them provides an explanation of P, then the explanatory argument provides better understanding of P. For instance, classical electrodynamics is unable to explain the stability of atoms. Quantum mechanics supplies an explanation. Therefore, quantum mechanics provides a better understanding of atoms than classical electrodynamics.

In general, an explanatory argument provides better understanding than a non-explanatory one. Given the received view, this claim might qualify as a conceptual truth. On my expressivist account, accepting this claim amounts to expressing acceptance of the following norm: we ought to intellectually prefer explanatory arguments over non-explanatory ones. Hence, the received view of understanding underwrites a rather obvious norm on comparative understanding. Moreover, one could endorse this norm while augmenting the received view, e.g. by holding that understanding-why is not exhausted by grasping explanations. Indeed, conceptualism purports to give such arguments against explanationism.

The norm that explanations provide better understanding than non-explanations has a rather natural extension. Based on how we commonly talk about explanations, 'explanation' appears to be a degreed notion. Given two explanatory arguments, one of them might provide a better explanation than the other (even if both are equally accurate). This observation motivates a further norm on comparative understanding: given two explanatory arguments, if one provides a better explanation, then it provides better understanding. If we accept the received view of understanding, then this norm takes on the air of a conceptual truth. Insofar as understanding amounts to grasping an explanation, grasping a better explanation leads to better understanding. Phrased as an explicit norm on intellectual preference, this yields the following: we ought to intellectually prefer better explanations.

I expect that these are two of the least controversial norms one could endorse on comparative understanding. To a large extent, Khalifa's (2017) account of understanding consists in spelling out the constraints that explanation places on comparative understanding. Defending a form of explanationism, Khalifa advocates two conditions for greater understanding: either, an agent grasps more explanatory information about the phenomenon, or their grasp of this information better resembles scientific knowledge (2017, p. 14). Framed explicitly as norms on intellectual preference, Khalifa's claims amount to the following: i) we ought to intellectually prefer grasping more explanatory information to less, and ii) we ought to intellectually prefer greater justification, resilience, safety, or other epistemic goods that contribute to scientific knowledge.

Due to the arguments in Section 4.2.2 and Chapter 3, I am doubtful that these norms capture all of the intellectually significant features of scientific understanding. This is because I believe we can have intellectually significant differences without explanatory differences (where the grasp of these differences equally resembles scientific knowledge). Hence, the extent to which we grasp explanatory information does not exhaust sources of better understanding.

The first family of norms, then, arises from constraints that explanation places upon better understanding. Of course, how to apply these norms depends very much on one's account of explanation, an exceedingly controversial topic within philosophy of science (and philosophy generally). Reasonable people frequently disagree about which of two explanatory arguments provides a better explanation; likewise for the question of whether an argument is even explanatory (as opposed to merely providing a derivation or justification for belief). Regardless of their preferred account of explanation, seemingly everyone can agree on the preceding two norms.[34]

4.6.2 Norms from number of EDRs

The second family of norms arises from differences in the number of EDRs or intellectually significant properties that a formulation provides. In some cases, one formulation extends another, more elementary formulation. The extension then underwrites strictly more epistemic dependence relations than the elementary formulation. The Wigner–Eckart theorem provides one such example: it extends an elementary approach to calculating matrix elements in atomic, nuclear, and molecular physics (Hunt 2021a). Likewise, Feynman diagrams extend an elementary approach to calculating scattering amplitudes in particle physics (see Section 3.7). In these cases, additional EDRs tell us that there is some information that we don't need to know in order to solve problems, e.g. to calculate particular physical quantities. Whereas the elementary formulations rely on this information to solve problems.

It is natural to think that the extended formulation is intellectually superior to the elementary one. After all, it contains the elementary formulation as a proper sub-plan for

[34]Philosophical disagreements about explanation typically concern what it takes for one fact to be explanatorily relevant for another. Elsewhere, I defend an expressivist account of explanation, based on expressivism about explanatory relevance. To judge that a fact is explanatorily relevant to answering a why question is to express an attitude of *being for being satisfied* by that fact as part of an answer.

solving problems. These observations motivate the following claim about comparative understanding: a formulation that provides strictly more epistemic dependence relations than another provides better understanding. Formulated explicitly as a norm, this yields the following: other things equal, one ought to intellectually prefer a formulation that provides strictly more epistemic dependence relations.

To avoid pedantic counterexamples, some qualifications are required. Not just any gain in epistemic dependence relations will do. We are only interested in those that actually contribute to solving problems about a particular set of phenomena. One could in principle adjoin arbitrarily many epistemic dependence relations that have nothing to do with the given phenomena. If an EDR does not contribute to a problem-solving plan, then it is irrelevant for that plan.

At first glance, the formulations of crystal field theory appear to fall into this special case where a more sophisticated formulation extends a more elementary formulation. The group theoretic approach builds on the elementary approach. It contains the elementary approach as a proper subset of its expressive means. The norm above then seems to license the claim that the group theoretic approach provides better understanding than the elementary approach. While this norm applies at least in the context of the eigenvalue problem, it does not apply in all contexts. For problems concerning splitting and degeneracy, we can restrict the group theoretic approach to a problem-solving plan that does not appeal to the elementary approach. We can then ask whether the elementary or the group theoretic approach provides a better understanding of the relevant phenomena. Within this context, neither approach is an extension of the other. Hence, we cannot appeal to a norm based on counting epistemic dependence relations. Instead, we need to consider norms from the third family, introduced below.

Unsurprisingly, it is often not the case that a more sophisticated approach builds upon a less sophisticated one. For instance, the group theoretic approach does not contain the non-group theoretic symmetry-based approach as a proper subset of its expressive means. Generically, compatible reformulations can bear complicated inferential relations to one another, with neither being an extension of the other. Each formulation might supply its own set of epistemic dependence relations, with different intellectually-significant properties. Figuring out which—if any—formulation provides better understanding requires comparing these different EDRs and resultant properties, weighting them to different degrees. For example, perhaps EDRs that unify are more epistemically valuable, *ceteris paribus*, than EDRs that merely provide a uniform treatment. These considerations lead us to the third family of norms, namely norms governing which kinds of EDRs one ought to intellectually-prefer.

Before moving on to this third, more interesting family of norms, it is worth pointing out a consequence of the preceding simple norm on counting EDRs. It answers the following kind of comparative question about understanding: together, do two formulations for solving the same problem provide better understanding than either one alone? Intuitively, the answer is often 'yes,' two formulations are often better than one. The simple norm on counting EDRs provides one way of underwriting this intuition, while providing principled reasons for cases where two formulations are *not* better than one.

Recall that in order for two formulations to be significantly different—as opposed to trivial notational variants—they must differ in at least one EDR. Typically then, two significant reformulations provide strictly more EDRs than either one on their own (an exception is the case where one formulation is an extension of the other). If the norm on counting EDRs is correct, then we ought to intellectually-prefer having more EDRs to fewer. Hence, we ought to intellectually-prefer having more significant reformulations to fewer. This entails that two significant reformulations provide better understanding than either alone (barring the extension case).

For instance, independently of whether or not the group theoretic account provides a better understanding than the elementary approach, the conjunction of these two reformulations provides a better understanding of crystal field theory than either formulation on its own. This follows simply from the fact that both reformulations together provide more epistemic dependence relations than either approach.[35] In contrast, in cases of trivial notational variants, both formulations provide the same EDRs. Hence, it is not the case that we ought to intellectually-prefer having more trivial notational variants. We do not understand a phenomenon better by knowing how to treat it using two different conventions, when those conventions do not lead to any epistemic differences in problem-solving plans.

[35]For example, the elementary approach shows that no knowledge of group representation theory is needed to solve the crystal field theory problem. It also shows that one does not need to know the symmetry of the coordination complex.

4.6.3 Norms on kinds of EDRs

This leaves the most interesting norms on comparative understanding. These norms pertain to EDRs and the intellectually significant properties that different EDRs can generate. Such properties include modularization, unification, uniformity of treatment, and manifestness. There are probably others that outstrip my limited experience.[36] For each of these properties, a formulation that instantiates it might provide a better understanding than a formulation that lacks the given property. We can schematize such comparative claims as follows: *ceteris paribus*, solving a problem with a formulation that possesses an intellectually significant property Q provides a better understanding than one that lacks Q. On my expressivist account, accepting this claim amounts to expressing acceptance of a set of norms that recommend intellectually-preferring this property.

At the first-order level then, we have to settle at least two questions. First, which EDRs and properties ought we intellectually-prefer? Second, how should we weight these EDRs or properties in cases where two formulations exhibit different ones? For instance, one formulation might modularize while another makes a property manifest. Other things equal, judging that the modular formulation provides better understanding than the manifest formulation would amount to expressing acceptance of a set of norms where we ought to intellectually-prefer modularization over manifestness.

The first question seems much easier to answer than the second. I am inclined to think that we ought to intellectually-prefer those epistemic dependence relations that give rise to intellectually significant properties. This is because I think we ought to intellectually-prefer such properties. Other things equal, it seems intellectually better for a formulation to have these properties than to lack them. Hence, I am inclined to endorse a system of norms that recommends intellectually-preferring the properties of modularization, unification, uniformity of treatment, manifestness, and plausibly others that I have left out. Developing formulations that exhibit these properties contributes to the clarification of epistemic structure. It clarifies what we need to know to solve problems. In Chapter 5, I argue that clarifying epistemic structure is a constitutive aim of science. Hence, I believe that developing formulations with these properties has final value for science (as opposed to merely instrumental value).

[36] Perhaps visualizability is an intellectually significant property, connected with the nature of geometric EDRs, as opposed to algebraic ones.

On the second question, regarding how to weight our intellectual preferences for these properties (or EDRs that give rise to them), I am not confident in any particular answer. An analogous question arises in the moral domain: how are we to weight competing moral goods when they come into conflict, such as respecting autonomy vs. promoting well-being? A utilitarian has a straightforward answer: do whatever best promotes utility for the greatest number in the long run. A similar answer is available here: perhaps some intellectual virtues are more likely to lead to more accurate theories in the long run. An epistemic utility theorist might then favor these virtues, whatever they are.

Thinking instrumentally in this way, it might be the case that science on the whole does best when different scientists subscribe to different weightings of intellectually significant properties. Different weightings might motivate scientists to pursue different projects, and one cannot know in advance which problem-solving strategies are most likely to bear fruit for currently unsolved problems. The case study methodology I pursue here can at best provide inductive grounds for privileging certain intellectually significant properties. Nevertheless, the question seems like it might not admit a decisive answer. It is similar to figuring out how best to weight various artistic virtues, so as to figure out which of two great works of art is aesthetically better. Such weightings might ultimately just be matters of taste.

4.7 Conclusion

I have argued that non-explanatory, intellectual differences arise even outside the context of theoretically equivalent reformulations. The three formulations of crystal field theory provide a paradigmatic example of theoretically *inequivalent* formulations that nevertheless pose a counterexample to explanationism. Moreover, I have shown how conceptualism provides an account of the intellectual significance of symmetry arguments in cases where they are not needed to explain phenomena. Section 4.2 illustrates the dramatic intellectual differences that symmetry arguments can provide. Group representation theory radically restructures how we understand the energy levels of coordination complexes. It does this by modularizing the crystal field theory problem into separately treatable subproblems, while unifying systems into symmetry-based families.

Sections 4.3 and 4.4 considered what seem to be the only two strategies available for defending explanationism—short of denying that compatible formulations lead to differences in understanding. By rebutting these strategies, I have shown that existing accounts of explanation face the burden of accommodating non-explanatory intellectual differences. One easy way to meet this burden is simply to renounce explanationism and adopt conceptualism. Conceptualism provides a general approach to interpreting the intellectual and methodological significance of reformulations. This includes mathematical reformulations in particular, which have recently sparked debates over the existence of non-causal explanations. Promisingly, conceptualism lets us interpret mathematized explanations while skirting seemingly insoluble metaphysical disputes. It focuses attention away from epistemically inaccessible features of scientific ontology and toward the manifestly accessible epistemic structure of problem-solving plans.

In the final two sections, I considered the question of what makes one formulation intellectually better than another. I argued that to answer this question in full generality, we need to appeal to the preferences of scientific agents. Section 4.5 proposed a novel expressivist account of the better-understanding-than relation. Comparative judgments of intellectual value express preferences for specific kinds of epistemic dependence relations or intellectually significant properties. Section 4.6 considered a number of first-order norms on intellectual preference, underwriting specific claims about comparative understanding. In Section 6.6, I will return to expressivism in order to develop a non-metaphysical account of fundamentality.

A Appendix

A.1 The elementary approach

The elementary approach solves the crystal field theory problem exclusively through perturbation theory. We begin with an initial Hamiltonian H^0 (with known eigenvalues and eigenfunctions) that characterizes the energy and dynamics of the unperturbed system, such as Ni^{2+}. We characterize the perturbation from the six water molecules by an operator H'. The sum of these two operators equals an approximate Hamiltonian, H, for nickel(II) hexahydrate: $H = H^0 + H'$. To approximate the eigenvalues of H, we first calculate the matrix elements of the perturbation operator H' by measuring the electrostatic potential. Calculating these matrix elements requires choosing a basis for the five unperturbed d-orbitals of Ni^{2+}, such as the five spherical harmonics, $Y^{2,m}$ ($m \in \{-2,-1,0,1,2\}$).[37] In this basis, we compute all twenty-five elements of the 5×5 matrix H' and use them to solve the secular equation of the perturbation operator:

$$determinant[H' - \lambda I] = 0 \tag{A.1}$$

where λI is a constant multiple of the identity matrix. The roots (i.e. zeros) of the secular equation equal the eigenvalues of H', which provide a first-order correction to the eigenvalues of H^0. The number and degeneracy of the distinct eigenvalues corresponds to the number of new energy levels and their degeneracies.

A.2 The non-group-theoretic approach

The non-group-theoretic approach begins by determining a general form for the electrostatic potential in terms of the symmetry of the coordination complex. Using Coulomb's law, we express the potential at an arbitrary point P as a sum of six contributing potentials, one from each of the ligands. After manipulating this expression using Legendre polynomials, we arrive at a tractable formula in Cartesian coordinates x, y, and z. This constitutes a *symmetry-based form for the potential* (Figgis and Hitchman 2000, p. 38):

$$V = \sum_{i=1}^{6} V_i = 6\frac{Ze^2}{a} + \frac{35Ze^2}{4a^5}(x^4 + y^4 + z^4 - \frac{3}{5}r^4) \tag{A.2}$$

Here, Z is the charge of the central metal ion, r is the distance from the point P to the central metal ion, and a is the distance between each ligand and the central metal ion.

Using equation (A.2) for the crystal field potential V, we then proceed as in the elementary approach. We calculate the matrix elements of the perturbation operator H' and solve the resulting secular equation for its roots. Since this derivation uses a symmetry-based expression for the potential, it applies not just to nickel(II) hexahydrate but to any coordination complex with octahedral symmetry instantiated in the same way. We find two distinct roots: $\lambda_1 = -\frac{2}{5}\Delta_O$ (three-fold degenerate) and $\lambda_2 = \frac{3}{5}\Delta_O$ (two-fold degenerate), expressed in terms of their energy difference Δ_O. The existence of two distinct roots entails that two new energy levels form. Since these roots are three-fold and two-fold

[37] These functions belong to the separable Hilbert space $L^2(\mathbb{R}^3)$ of square-integrable complex-valued functions defined over Euclidean three-space $\mathbb{R}^3$. For details, see Cornwell (1984, Appendix B).

degenerate, so are the resulting energy levels (Dunn et al. 1965, p. 16).

A.3 The group-theoretic approach

To apply group theory, we first identify the symmetry group G^0 of the unperturbed Ni^{2+} metal ion.[38] This is the set of transformations that leave the initial Hamiltonian, H^0, invariant. An unperturbed metal ion possesses spherical symmetry, so H^0 is invariant under any rotation around any axis passing through the centre of Ni^{2+}. This uncountably infinite set of rotations constitutes the *pure rotation group* $SO(3)$, (i.e. the special orthogonal group in three dimensions).[39]

Next, we identify the symmetry group G of the final, perturbed state—in this case nickel(II) hexahydrate. In Figure 6, Ni^{2+} sits at the centre of an octahedron, surrounded by a water ligand at each vertex. The symmetry operations that leave this coordination complex's Hamiltonian, H, invariant are the twenty-four operations of the octahedral group, O. These comprise a variety of $90°, 120°$, and $180°$ rotations through various axes passing through the octahedron's centre.

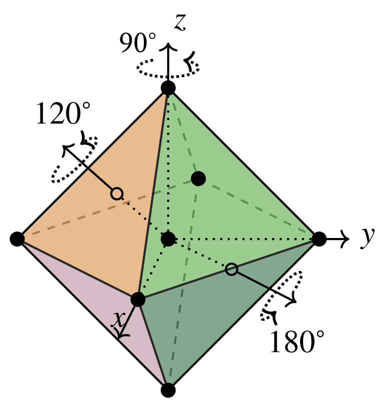


Figure 6: Some symmetry operations of an octahedron

These symmetry transformations form five distinct 'conjugacy classes': $(1)\,\mathbf{360°}$, $(8)\,\mathbf{120°}$, $(6)\,\mathbf{180°'}$, $(3)\,\mathbf{180°}$, and $(6)\,\mathbf{90°}$, corresponding to distinct kinds of rotations. Here, '$(1)\,\mathbf{360°}$' indicates a conjugacy class consisting of one 360 degree rotation, namely the

[38]A *group* is a set equipped with a closed, invertible, and associative binary operation, containing an identity element.

[39]To simplify the exposition, I neglect inversion transformations in both the initial and final symmetry groups.

identity element of the group.[40]

To extract information about the energy levels from these symmetry groups, we move from group theory to group representation theory. This involves constructing 'matrix representations' for the symmetry groups of interest.[41] To form a matrix representation, we map each geometrical symmetry transformation B to an invertible matrix $\rho(B)$. For finite groups, such as the octahedral group, each representation can be decomposed into a finite number of 'irreducible representations.'[42] Since these irreducible representations cannot be decomposed further, they function as the basic building blocks of all other representations. For instance, the octahedral group has five irreducible representations, labeled A_1, A_2, E, T_1, and T_2.[43]

Irreducible representations express important symmetry properties of physical systems. In particular, the irreducible representations of a system's symmetry group label the Hamiltonian's eigenvalues, i.e. the energy of each orbital. This means that each distinct energy level (each distinct eigenspace) corresponds to an irreducible representation.[44] For example, an irreducible representation '$\Gamma_{rot}^{(2)}$' of $SO(3)$ labels the five-fold degenerate d-orbitals of Ni^{2+}. This correspondence between eigenspaces and irreducible representations allows us to derive facts about energy levels by considering relationships between representations.

For many applications, it is unnecessary to determine explicit matrix representations for each irreducible representation (illustrating another epistemic dependence relation). Instead, we can often rely on group characters. For a given irreducible representation, a 'character' is the trace of a matrix from that representation (the 'trace' is the sum of the elements along the principal diagonal). Since matrix traces are invariant under changes in

[40]A *conjugacy class* is a collection of operations that is invariant under conjugation: letting A be a member of the conjugacy class and X any member of the group, the combination XAX^{-1} is also a member of the conjugacy class.

[41]A matrix representation is a group homomorphism ρ from the group of interest to the group of invertible linear transformations over a vector space V (i.e. the general linear group $GL(V)$). Requiring this map to be a 'group homomorphism' means that the matrix representatives must compose under matrix multiplication in the same way as the symmetry transformations do: $\rho(AB) = \rho(A)\rho(B)$.

[42]A representation is *irreducible* if there is no proper subspace of basis vectors left invariant by the transformations of the symmetry group, i.e. the vector space for the representation contains no smaller, nontrivial invariant subspaces. Otherwise, a representation is *reducible*.

[43]The capital letters correspond to the dimensionality of the irreducible representation, with 'A,' 'E,' and 'T' corresponding to one, two, and three dimensional representations, respectively.

[44]In general, each d-dimensional eigenspace provides a basis for a d-dimensional representation of the Hamiltonian's symmetry group. In cases of accidental degeneracy, this representation is reducible, rather than irreducible.

basis, characters are invariants of an irreducible representation, meaning that they do not depend on the basis chosen for the representation. Thus, each irreducible representation has a well-defined set of characters. Furthermore, since the trace of a matrix is invariant under conjugation, members of the same conjugacy class have the same trace, and thus the same characters. As a result, we can organize the characters in a table, where the rows label the irreducible representations and the columns label the conjugacy classes. The character table for the octahedral group is shown in Table A.1.

Table A.1: Character table for the octahedral group

$\mathbf{O}$	(1) $\mathbf{360}^\circ$	(8) $\mathbf{120}^\circ$	(3) $\mathbf{180}^\circ$	(6) $\mathbf{180}^{\circ\prime}$	(6) $\mathbf{90}^\circ$	Good basis functions
$\mathbf{A_1}$	1	1	1	1	1	$x^2+y^2+z^2$
$\mathbf{A_2}$	1	1	1	-1	-1	
$\mathbf{E}$	2	-1	2	0	0	$(2z^2-x^2-y^2, x^2-y^2)$
$\mathbf{T_1}$	3	0	-1	-1	1	
$\mathbf{T_2}$	3	0	-1	1	-1	(xz, yz, xy)
$\Gamma_{rot}^{(2)}$	5	-1	1	1	-1	

Affixed to the bottom of Table A.1 are the characters of the representation $\Gamma_{rot}^{(2)}$ for each conjugacy class of the octahedral group. These characters follow from a general equation for the character of a rotation through α radians for an irreducible representation characterized by angular momentum ℓ (Cotton 1990, p. 261):

$$\chi^{\ell}(\alpha) = \frac{sin[(\ell+\frac{1}{2})\alpha]}{sin[\alpha/2]} \tag{A.3}$$

For instance, the character of the $\Gamma_{rot}^{(2)}$ matrix representatives for 120° rotations is -1, obtained by substituting '$\frac{2\pi}{3}$ radians' for α and '2' for ℓ (since we are dealing with d-orbitals, which have an orbital angular momentum of two).

When we perturb the spherical symmetry by surrounding the metal ion with ligands, we break its spherical symmetry into octahedral symmetry. Consequently, irreducible representations from the octahedral group now label the energy levels of the system. Therefore, the irreducible representation $\Gamma_{rot}^{(2)}$ that labels the initially degenerate energy levels decomposes into a direct sum of irreducible representations Γ^r from the octahedral group:

$$\Gamma_{rot}^{(2)} \approx \sum_r \oplus n_r \Gamma^r \tag{A.4}$$

where the multiplicity, $n_r \in \mathbb{N}$, indicates the number of times that the irreducible representation Γ^r occurs in this decomposition.

Each distinct Γ^r in the decomposition (A.4) labels an eigenspace of nickel(II) hexahydrate. Thus, the number of distinct Γ^r in the decomposition provides the splitting, i.e. the number of distinct eigenspaces, and the dimension of Γ^r equals the degeneracy of the corresponding eigenspace. Hence, the decomposition of $\Gamma^{(2)}_{rot}$ determines the splitting and degeneracy.

To determine this decomposition, we do not need to know explicit matrix representations for the irreducible representations. Instead, it suffices to use a *character decomposition formula*:

$$n_r = (1/g)\sum_k N_k \chi^r(T)^* \chi_{rot}(T) \tag{A.5}$$

This equation provides a general relation for decomposing a reducible representation into a sum of irreducible representations of a finite group (Cornwell 1984, p. 85). As before, n_r denotes the multiplicity of the r-th irreducible representation in the decomposition of $\Gamma^{(2)}_{rot}$. g denotes the cardinality of the symmetry group of the coordination complex, in this case the octahedral group (which has 24 elements). The sum is taken over each conjugacy class (i.e. column) of the character table, indexed by k. N_k denotes the number of symmetry operations in the k-th conjugacy class. In Table A.1, N_k corresponds to the number preceding the type of symmetry operation at the top of the table. $\chi^r(T)^*$ denotes the complex conjugate of the character of a symmetry operation T in the k-th class for the r-th irreducible representation. Since the characters for the octahedral group are real, these are simply the characters in Table A.1. Likewise, $\chi_{rot}(T)$ denotes the character of the same symmetry operation for the representation $\Gamma^{(2)}_{rot}$.

To illustrate the use of Equation A.5, we can compute the multiplicities n_r in the decomposition of the representation $\Gamma^{(2)}_{rot}$. In each equation below, the three factors in the k-th summand all come from the k-th column of Table A.1. The first number in each summand is the class size N_k from the top row. The second number is $\chi^r(T)$ from the row of the irreducible representation. The third number is $\chi_{rot}(T)$ from the table's bottom row.

$$
\begin{aligned}
n_{A_1} &= \frac{1}{24}[(1)(1)(5)+(8)(1)(-1)+(3)(1)(1)+(6)(1)(1)+(6)(1)(-1)] = 0 \\
n_{A_2} &= \frac{1}{24}[(1)(1)(5)+(8)(1)(-1)+(3)(1)(1)+(6)(-1)(1)+(6)(-1)(-1)] = 0 \\
n_{E} &= \frac{1}{24}[(1)(2)(5)+(8)(-1)(-1)+(3)(2)(1)+(6)(0)(1)+(6)(0)(-1)] = 1 \\
n_{T_1} &= \frac{1}{24}[(1)(3)(5)+(8)(0)(-1)+(3)(-1)(1)+(6)(-1)(1)+(6)(1)(-1)] = 0 \\
n_{T_2} &= \frac{1}{24}[(1)(3)(5)+(8)(0)(-1)+(3)(-1)(1)+(6)(1)(1)+(6)(-1)(-1)] = 1
\end{aligned}
\tag{A.6}
$$

This calculation shows that the only irreducible representations that occur in the decomposition of $\Gamma_{rot}^{(2)}$ are E and T_2. Thus, the fivefold degenerate d-orbitals split into two new energy levels, with symmetry type E and T_2, respectively. This solves the splitting problem. It also solves the degeneracy problem. In Table A.1, the character of the identity transformation (found in the column under 360°) equals the dimension of the corresponding irreducible representation. Thus, we see that the E irreducible representation is two-dimensional, while the T_2 irreducible representation is three-dimensional. Recalling that the dimension of an irreducible representation equals the dimension of the corresponding eigenspace, we see that the resulting energy levels are two-fold and three-fold degenerate, respectively. Finally, through the method of projection operators, representation theory allows us to determine good basis functions that diagonalize the perturbation operator, H' (Cornwell 1984, pp. 92–8). These are listed in the final column of the character table. They help modularize the eigenvalue problem into separate calculations for each eigenvalue.

Chapter 5:

Reformulation as a Constitutive Aim of Science

5.1 Introduction

In developing constructive empiricism, van Fraassen (1980) framed the debate over scientific realism as a disagreement about the aims of science. Scientific realism takes the aim of science to be approximate truth, including claims about physical unobservables.[1] In contrast, constructive empiricism defends a weaker aim for science: empirical adequacy. Empirical adequacy requires truth only about observable states of affairs. Nevertheless, as Rosen (1994) pointed out, the foundational status of these "aims of science" is not clear. Are they descriptive claims about scientific sociology? Prescriptive claims about what scientists ought to aim at, independently of their current goals? Or are they merely a useful way of talking, perhaps a kind of fictional story about science?

I will analyze the aims of science by focusing on those aims that are constitutive of scientific activity. Constitutive aims define the *minimal* criteria of success, i.e. the criteria that must be met for scientific activity to succeed at all. Unlike success criteria in general, constitutive features are often hidden from participants. For instance, native language speakers are often unaware of the grammatical rules they follow. I will argue that one constitutive aim of science is the ability to solve all physically possible empirical problems. I will call this aim *problem-solving adequacy*. My proposal is similar to Laudan's, who argues that "the aim of science is to secure theories with a high problem-solving effectiveness" (1996, p. 78). Of course, realists and antirealists disagree about the minimal success criteria for solving a scientific problem. Whereas realists posit truth, construc-

[1] Philosophers commonly speak about the aims of science, especially physics. See, for instance, Loewer (2007, pp. 319, 322, 326), who in passing distinguishes the aims of physics from the aims of metaphysics, or Potochnik (2017). Earman and Roberts "presuppose that science does aim to discover laws (among lots of other things)" (2005, p. 254). Section 5.8 considers objections to speaking about aims of science.

tive empiricists posit empirical adequacy. Because my account enriches both realism and antirealism, I will remain neutral on this contentious debate here.[2]

To argue that problem-solving adequacy is a constitutive aim of science, I focus on a ubiquitous aspect of scientific practice: reformulating scientific theories and problem-solving procedures. Unamended, both realism and antirealism leave mysterious why scientists go to such great lengths to clarify which epistemic resources they need to solve problems, even when this clarification does not make their theory any more empirically adequate or true. I will argue that beyond seeking truth or empirical adequacy, scientists ought to clarify the epistemic structure of their theories. This constitutive aim arises from the need to prepare for any possible problem-solving context that could arise. My argument shows that reformulations are not merely instrumentally valuable for either truth or empirical adequacy. Instead, reformulations are valuable for their own sake, since they are constitutive of clarifying epistemic structure, which is itself a constitutive aim of science.

The argument of this chapter rebuts a worry raised in Chapter 2. There, I wondered whether the value of reformulating might ultimately just be instrumental. If so, then conceptualism would itself be a kind of instrumentalism about reformulations (Section 2.4). Here, I show that conceptualism does not reduce to instrumentalism. Reformulations possess a kind of final epistemic value that goes beyond their instrumental value.

5.2 Constitutive Aims

Empiricism and Scientific Aims

Constructive empiricism and scientific realism posit the same truth conditions for a scientific theory: a theory is true only if it matches reality at all length scales.[3] Furthermore, constructive empiricists agree with realists about the meaningfulness of claims about physical unobservables (even ones that are in-principle empirically inaccessible). For these reasons, constructive empiricism is not a variety of verificationism. Instead,

[2]Elsewhere, I argue that empirical adequacy, rather than truth, is the correct minimal success criterion for solving a physical problem. Since agents lack any competence for perceiving unobservables, the constitutive aim of perception is accuracy about observables. Developing this argument would take us too far afield here.

[3]Traditionally, these positions rely on a correspondence theory of truth.

constructive empiricism restricts the aim of science from truth at all length scales to truth about observables. Aiming for truth about observables is equivalent to aiming for empirical adequacy.

More precisely, a theory is empirically adequate provided that it accurately describes and predicts all true claims about observables.[4] Colloquially, an empirically adequate theory 'saves the phenomena.' Although less demanding than achieving truth, empirical adequacy remains a difficult aim to satisfy. A theory is empirically adequate only if it accurately represents the truth about observables past, present, and future, including not just what will be observed, but what could be observed. As I discuss in Section 5.4, empirical adequacy—like truth—is best understood as a *futuristic* aim: we will plausibly never obtain an empirically adequate theory. Rather, we approach greater empirical adequacy over time.

Rosen (1994) points out that as a sociological matter, many (if not most) scientists are aiming for more than empirical adequacy. Hence, for constructive empiricism to be plausible, it must distinguish the aims of science from the aims of individual scientists, even their aggregate aims. Indeed, van Fraassen contends that the aims of science are not the same as the aims of most or even all scientists (1994, p. 185). This generates a distinction between intentionality at the level of agents vs. intentionality at the level of the activity of science. Intentionality at the agent-level includes the intentions, opinions, and motivations of individual scientists. Intentionality at the activity-level abstracts away from these individual intentions. The aims of science exist at the activity-level.

To insulate the aims of science from the intentions of individual scientists, van Fraassen characterizes aims as criteria of success (1980, p. 8). As he notes, criteria of success often exist independently of the personal aims or motivations for engaging in a given activity. As a guiding analogy, both van Fraassen and Rosen consider the game of chess. One aim of chess is to checkmate your opponent. This aim serves as a shared criterion of success, independently of the particular reasons why anyone is playing chess. It is in this sense that van Fraassen claims that the "criterion of success in actual practice is empirical adequacy," regardless of scientists' individual aims, motives, or intentions

[4]On the semantic (model-theoretic) view of theories, this requires that the empirical substructure of the theory's models is isomorphic to observable reality (van Fraassen 1980, p. 64). However, I intend to remain neutral on the debate between syntactic vs. semantic approaches to scientific theories, especially given recent arguments that they are not competitors (Halvorson 2016; Lutz 2017).

(1994, p. 182). It is at least plausible that empirical adequacy could be an aim of science at the activity-level without being an explicit aim at the agent-level.

Nevertheless, van Fraassen's proposal is too weak to exclude truth from being an aim of science. If the aims of science are merely criteria of success, there is no reason why truth cannot be an aim as well. For surely, obtaining truth about unobservables would be a great success, in addition to empirical adequacy. Hence, interpreting the aims of science as criteria of success does not favor constructive empiricism over scientific realism. Similarly in chess, a player would demonstrate supreme mastery if they were able to not only checkmate their opponent but also capture all of their pieces along the way (or if they were able to checkmate their opponent as quickly as possible). Noting that checkmating your opponent is a criterion for success does not exclude other criteria for success. Likewise, recognizing that empirical adequacy is a criterion for success does not exclude either aiming at the truth or aiming at fundamental structure.

Van Fraassen's proposal faces a further, related challenge. Rosen wonders why the aims of science—defined as criteria of success—are not themselves "constituted by the conscious understandings of the participants" (1994, p. 146). Why isn't it the case that scientists' own understanding of science determines what qualifies as success? In response, van Fraassen grants that the conscious understandings of scientists define the activity 'science,' but that nevertheless this "does not mean what all the participants say they are doing is what they are doing" (1994, p. 186). This response grants the possibility that all scientists aim at truth, without truth being a criterion of success for science. But this response is implausible for success criteria in general. If the majority of scientists decided that the gold standard of experimental design requires insulating all measurement devices from Wi-Fi signals, this would generate a success criterion. It is therefore implausible that success criteria in general are independent from the conscious understandings of scientists.

Aims that are Constitutive

Fortunately for the aspiring empiricist, van Fraassen's response stands a chance when restricted to a more narrow class of success criteria. I will call these *constitutive aims* of an activity. They define *minimal* criteria of success. In chess, a constitutive aim is to checkmate your opponent: this is not merely a criterion for success; it is a minimal criterion.[5] In contrast, capturing all of your opponents' pieces—while no doubt grounds for success—is not constitutive of this success. Even if all chess players and chess federations agreed tomorrow that winners would receive twice the prize money for capturing all pieces, this would not make capturing all pieces into a constitutive aim. Capturing all pieces would still be unnecessary for winning a game of chess.

To understand what it takes for a criterion to be minimal, I invoke the notion of a constitutive feature. 'Constitutive aims' are then *the constitutive criteria for success*, determined by the activity itself. A *constitutive feature* is a feature something has in virtue of being itself. A thing's constitutive features are its conceptually essential or indispensable aspects. For instance, constitutive of the bishop piece in chess is moving along diagonals. Constitutive features, including constitutive rules, stand in contrast to arbitrary conventions, such as designating a particular piece as 'the king' (Searle 1995, p. 49). Instead, they are connected to the very possibility of the object or activity. As Max Black notes, "Chess played for the sake of losing is not chess but some other game" (1970, p. 159).[6]

In his accounts of perception, action, and knowledge, Sosa (2015) invokes a similar notion of a constitutive aim: a performance succeeds provided that it attains its constitutive aim. Moreover, attaining a constitutive aim does not preclude a performance from succeeding "even more fully" (2015, p. 14). It is in this sense that I take constitutive aims to define criteria for minimal success.[7]

Unlike aims in general, constitutive aims have the features that van Fraassen's account requires. Constitutive features, rules, or aims are often not manifest or perspicuous to participants. For instance, native speakers of English are often unaware of its grammatical rules, even rules that they routinely follow. Their lack of awareness is not evidence that there are no grammatical rules.[8] Moreover, English speakers could collectively decide to institute new rules, but this would not necessarily affect constitutive rules of their

[5]Similarly, we can understand a constitutive aim of an individual chess move as being "to do what will best help the player towards winning, or at least toward averting immediate defeat" (Sosa 2015, p. 126).

[6]We also commonly talk about constitutive norms. In her discussion of Peirce, Misak claims that "it is a constitutive norm of belief that a belief is responsive to the evidence and argument for or against it" (2013, p. 35).

[7]Sosa prefers to say that agents *succeed fully* when they attain a constitutive aim, but I worry this leads to grammatical confusion: in many contexts, it seems strange to say that "you can succeed fully even if you might have succeeded even more fully" (Sosa 2015, p. 14). The notion of minimal success preserves the relevant distinction while avoiding this infelicity.

[8]For a similar discussion in the context of semantic rules, see Thomasson (2020, p. 65).

language. As another example, many people are unaware that the value of money is a social construct, but this does not make money's value any less socially constructed. People might even believe money is valuable due to holding a false theory, such as that it must be "backed by gold" (Searle 1995, p. 47). The upshot is that scientists may very well be unaware of or even mistaken about the constitutive aims of science, without their lack of understanding affecting these aims. Focusing on constitutive aims thereby clarifies why aims-talk is not inherently sociological. Constitutive aims support a distinction between the intentional aspects of science vs. the intentions or opinions of individual scientists.

In the next section, I propose a method for identifying constitutive aims of science. Before that, two points of clarification are in order. First, the constitutive aims of an activity need not be sufficient for defining that activity. Multiple different games might all have a constitutive aim of checkmating opponents. For this reason, the constitutive aims that I defend for science do not demarcate science from either pseudoscience or non-science. If they are honest investigators, cryptozoologists searching for Bigfoot are bound by empirical adequacy. Likewise, ordinary problem-solving (such as figuring out how to cook pasta) involves considering possible scenarios and striving for empirical adequacy. But we wouldn't count cooking pasta as science, unless it were investigated sufficiently systematically![9]

Second, science is typically a community activity. Hence, the normative labor of constitutive aims is distributed: scientific agents ought to desire that members of their epistemic community are collectively working toward satisfying all of the constitutive aims of science. Each individual agent does not shoulder the normative weight of these obligations themselves. For whatever reasons, there may be particular tasks that individual agents abhor. They simply ought to desire that some magnanimous colleague will someday pick up the slack. The constitutive aims of science thereby generate *pro tanto* reasons for action. Such reasons can be outweighed at the individual level by other considerations, but never canceled.

[9] I thank Angela Sun and Sumeet Patwardhan for raising these demarcation issues.

5.3 Scientific Planning as Problem-Solving

Isolating constitutive aims of science requires identifying constitutive features of scientific activity. I will argue that one relevant constitutive feature is *scientific planning*, consisting of empirical problem-solving. Essential to scientific activity is answering questions about the physical world. As a method for answering empirical questions, science takes many forms: describing the past (e.g. archaeology), predicting the future (e.g. meteorology), building devices, synthesizing substances, etc. In different ways, each of these scientific activities contributes different methods for planning about how to interact with the world. They are methods for planning how to solve empirical problems, i.e. problems about the physical world. We can therefore understand one constitutive activity of science as being a form of planning. Scientists are planning how to solve problems, which amounts to planning how to answer their questions about the physical world. Even in the historical sciences, empirical problem-solving involves planning how to gather and assess evidence about the past.

Given this constitutive activity, we can determine its constitutive goals. Doing so will take us closer to isolating minimal success criteria for science. Beginning with local instances of scientific planning, we simply need to determine the local constitutive goals. Since this planning amounts to trying to solve an empirical problem, the constitutive goal is evidently to *solve the problem*. The foundational question is thereby to determine the minimal success criteria for solving an empirical problem: What makes a putative answer to a scientific question minimally successful? Answering this question helps identify local constitutive aims of science. We can then generalize from these local aims to determine global constitutive aims of science.

Following Goldman (1986, pp. 126–127), we can always define a problem in terms of asking a question, and we can define a solution in terms of an answer to that question. Schematically, an agent S has a scientific problem Q if and only if Q is a question about the physical world and agent S wants to have a successful answer to Q. Different criteria for what constitutes a "successful answer" lead to different positions in the realism vs. antirealism debate. Whereas Goldman proposes that a proposition B must be true in order to be a successful answer to the question, a constructive empiricist will settle for an empirically adequate answer. Remaining neutral between realism and antirealism, I will say that a proposition B is a solution/answer to a problem/question Q if and only if (i) B is a potential answer to Q and (ii) B satisfies the minimal criteria for success.

To remain neutral between truth vs. empirical adequacy, I will refer to the relevant constitutive aim of science as *planning adequacy*. Achieving planning adequacy would amount to being able to solve every scientific problem. Since the constitutive aim of other disciplines is presumably a form of planning adequacy as well, we should technically disambiguate *physical problem-solving adequacy* from other forms of planning adequacy (e.g. mathematical, logical, philosophical, etc.). For convenience, I will refer to 'physical problem-solving adequacy' as 'planning adequacy,' leaving it implicit that the focus here is on science. As I will argue in Section 5.5, neither truth nor empirical adequacy is actually sufficient for planning adequacy (at least not for beings like us). There is a further, subsidiary, constitutive aim necessary for achieving planning adequacy. Of course, the notion of 'adequacy' is ambiguous along various dimensions. What does it mean to 'really' solve a problem, and what does it mean to solve 'every' scientific problem? I turn to these issues next.

5.4 Ideal vs. Non-ideal Constitutive Aims

The minimal success criteria vary depending on whether we are talking about non-ideal vs. ideal scientific practice. There are at least two different dimensions of idealization. First, we can focus on obtaining exact—as opposed to approximate—solutions to scientific problems. This generates a distinction between exact vs. approximate empirical adequacy, which is simply a special case of the well-trod distinction between exact vs. approximate truth. Second, we can focus on all possible empirical problems, as opposed to the problems that scientific agents will actually face over their history. This distinction between solving all-possible vs. actual problems is a variety of modal completeness. Our knowledge about empirical problem-solving varies from radically incomplete at the beginning of science to maximally complete at the 'end of science.'

Both of these ideals—exactness and completeness—are *futuristic*: they will plausibly never be realized in finite time by any actual scientific community. What makes them futuristic is that they aim at "some ideally improved descendent that is never expected to actually exist, but that would result if the process of improvement went on forever"

(Lewis 1984, p. 230). In this way, *ideal* constitutive aims define the minimal criteria for success in the 'end of science.' In contrast, *non-ideal* constitutive aims define the minimal criteria for actual science, up to the demands that we require in practice and the problems that we actually face.[10]

The notion of 'adequacy' in 'empirical adequacy' is ambiguous along both dimensions of idealization. When it comes to solving actual problems non-ideally, it suffices for a solution to be approximate. Exact empirical adequacy (or truth) is unnecessary for practical purposes. As van Fraassen notes, "empirical adequacy is stronger than what counts as success *in practice*" (1995, pp. 157, 144). At best, scientists achieve various grades of approximate empirical adequacy, namely, an approximate fit between the predictions of their theory or solution procedure and the data obtained. At least due to measurement limitations, exactness is unattainable. Non-ideally, science aims at either approximate empirical adequacy or approximate truth. For this reason, scientific realism does not face a special problem of making sense of approximate truth that constructive empiricism avoids. Approximate empirical adequacy just is a special case of approximate truth: it is approximate truth about observables.

Restricted to individual problems, truth and empirical adequacy are both local criteria for success. They become global criteria when we consider wider classes of problems or problem-types. Beyond having (approximately) true or empirically adequate solutions to particular problems, scientists should ideally be able to solve any solvable problem they might encounter. This aim generates the second dimension of idealization: problem-solving completeness. Whereas non-ideal planning adequacy requires planning for all actual problems that will be encountered, ideal planning adequacy requires planning for all possible problems that could be encountered.

In the context of empirical problem-solving, the relevant notion of 'possible' is physically possible.[11] Whether a scientist ought to be able to solve a problem depends on whether that problem concerns a physically possible process, either exactly or approximately. Classical mechanics, for instance, typically concerns processes that are approximately physically possible (since in a non-classical world, an object can at best be ap-

[10]In passing, Davidson refers to a "constitutive ideal of rationality" (1980 [1970], p. 223). My contention here is that problem-solving adequacy is a constitutive ideal of science.

[11]Similarly, mathematics concerns at least logical possibility, while metaphysics focuses on metaphysically possible problem-solving. The local success criteria for solving a metaphysical problem are perhaps determined by the norms governing imagination.

proximately classical). Of course, we often do not know in advance which processes are physically possible. Hence, when inquiring, scientists often consider epistemic possibilities that turn out to be physically impossible. When scientists try to solve problems concerning non-physically possible scenarios, I take this to be instrumentally valuable for solving problems that are physically possible. Nonetheless, solving these non-physically possible problems is not a constitutive aim of science, even if this activity has independent epistemic value.

Here, I intend to remain neutral on how best to understand physical possibility. No doubt, different scientific realists have different preferred conceptions of physical possibility. Likewise, it is an open question which account of possibility is best suited for empiricism.[12] The specific content of scientific planning adequacy requires choosing an account of physical possibility, but the schematic conception remains the same.

My conception of planning adequacy is inspired by Gibbard's notion of a *hyperplan*.[13] Gibbard defines a hyperplan as a maximal contingency plan; it is a plan covering every possible circumstance (2003, p. 54). Clearly, no finite scientific agent ever needs a hyperplan to successfully solve the problems they actually encounter. Nevertheless, in the limit of successful scientific inquiry, we ought to develop hyperplans for all scientific problem-types. Even at the more local level of solving a given problem-type, scientists should ideally have a hyperplan for solving this problem-type.

To summarize: in the limit of a successful science, an epistemic community will have a plan covering all physically possible scenarios of interacting with the world. They will thereby attain *scientific planning adequacy*, i.e. knowledge of how to solve any (physically) possible physical problem. If we take empirical adequacy as the constitutive aim of problem-solving, then these agents will know all truths about observables, attaining what we might call *observational omniscience*. If instead we take truth as the minimal local success criterion, then agents in the end of science would know all physical truths and would thereby be *physically omniscient*. If a being attains physical omniscience, further scien-

[12]One strategy I pursue elsewhere is to combine an expressivist account of counterfactuals with Lange's (2009a) account of meta-laws, where counterfactuals are the truthmakers for physical necessity. This yields an empiricist-friendly account of meta-laws. I believe that the best way to interpret this account is as specifying the norms that govern our talk of laws and counterfactual planning, leading to a normativist approach to physical modality similar to Thomasson's (2020) approach to metaphysical modality.

[13]For his own inspiration, Gibbard cites Savage's (1954) discussion of big vs. small worlds in the context of decision theory.

tific inquiry becomes impossible. At most, they could 'inquire' in the sense of retrieving knowledge from their memory. But they could never inquire further about the physical world, since they already know everything about it.[14] As a result of their omniscience, they would have no need for scientific planning or problem-solving. All problems would already lay solved.

Considering the awesome power of a true scientific theory, one might worry that the notion of planning adequacy is redundant. Given either a true theory or an empirically adequate theory, wouldn't we *ipso facto* have achieved realist or empiricist planning adequacy, respectively? For recall that the notions of truth and empirical adequacy are futuristic: in the limit, they already purport to cover all past, present, and future states of affairs. Why then do we need a separate notion of planning adequacy? In the next section, I argue that for logically-omniscient beings, the notion of planning adequacy is genuinely redundant. However, for logically-imperfect beings like ourselves, neither empirical adequacy nor truth is sufficient. To achieve planning adequacy, beings like us must clarify the epistemic structure of our theories and problem-solving procedures. This, in turn, requires that we reformulate our scientific theories.

5.5 Aiming for Epistemic Suitability

Although an omniscient agent could not inquire about the world, mere *logical* omniscience does not preclude inquiry. Assuming that the state of the world does not follow from logic alone, a logically-omniscient being still has much to learn. If they were to arrive at an empirically adequate theory, then they would have reached the end of scientific inquiry, according to constructive empiricism. When coupled with logical omniscience, an empirically adequate theory automatically grants empiricist planning adequacy. Empirical adequacy entails that for any possible physical scenario, the theory has a model whose empirical substructure matches the observable phenomena. A logically-omniscient agent would immediately know how to apply such a theory to save the phenomena. As I argue below, this is because they already have full knowledge of the theory's epistemic

[14] As Woodard (forthcoming) argues, agents can genuinely check claims they already know. Hence, by 'omniscient,' I technically mean an agent who not only knows everything, but is also certain of this knowledge. Perhaps a physically omniscient agent could inquire further into the fundamental structure of reality, e.g. about how to state truths in a perfectly natural language. Whether these are further facts about the physical world depends on the relationship between metaphysics and science.

structure. Hence, for logically-omniscient beings, an empirically adequate theory is already sufficient for possessing a scientific hyperplan. Logically-omniscient agents with an empirically adequate theory are thereby guaranteed to be minimally successful when it comes to solving all possible scientific problems. They possess a plan that predicts and accurately describes all observable parts of the world. *Mutatis mutandis*, the same remarks apply to a logically-omniscient agent with a true theory. They are guaranteed to be minimally successful according to scientific realism.

However, most agents are presumably logically imperfect, rather than logically omniscient. Logically-imperfect agents have a real need to struggle through the derivations of what entails what, or of whether a given fact is necessary for another. They are not disposed to immediately know the logical connections between their concepts. For these agents, neither empirical adequacy nor truth is sufficient for planning adequacy. Instead, they have a further constitutive need to clarify the epistemic structure of their theories (insofar as they are interested in scientific problem-solving at all). To show that this need gives rise to a further constitutive aim, I will assume that these agents already have a true theory (so *a fortiori*, an empirically adequate theory as well).[15] This will be the case that humans find themselves in if we ever arrive at a true scientific theory, since we will never be logically omniscient (at least on account of our finite minds).

As argued in Section 5.3, a constitutive activity of science is empirical problem-solving. Successful problem-solving requires more than a true theory: it requires being able to apply the theory.[16] For logically-omniscient beings, truth alone (or empirical adequacy) suffices for applicability. This is not so for logically-imperfect beings. At one extreme, a true theory would be mostly useless for an agent who could not reason at all. To borrow an analogy from computability theory, having a true theory is akin to knowing a class of decidable problems. Agents who know this will know what they can and cannot do with their theory, in principle. However, this is a far cry from actually knowing a decision procedure, i.e. possessing a method for going from 'decidable' to 'decided.' Successful planning requires the latter, not merely the former.

To solve problems, agents need to figure out what suffices for a solution. This requires

[15] We could even assume that these agents know the logical relationships between their concepts but not that these are the logical relationships. Their thinking about these relationships could be (temporarily) caught in the musings of Lewis Carroll's tortoise.

[16] In a similar vein, Loewer notes that "the information in a theory needs to be extractable in a way that connects with the problems and matters that are of scientific interest" (2007, p. 325).

formulating the theory such that it provides a usable problem-solving procedure, such as an algorithm. For non-futuristic problem-solving, agents simply need to know a set of sufficient conditions for solving the problems they will actually face. In the ideal of futuristic problem-solving, agents need to be able to solve any possible problem they might encounter. This requires figuring out what they need to know to solve problems of any given problem-type. Overall then, scientific agents ought to determine what they need to know and what suffices to know in order to solve any possible scientific problem. These necessary and sufficient conditions for problem-solving are precisely what I have been calling *epistemic dependence relations* (EDRs). Hence, we see that clarifying the epistemic structure of a theory amounts to determining EDRs, i.e. epistemic relations relevant for problem-solving.

These considerations show that for logically-imperfect agents, there is a further constitutive aim of science, beyond truth or empirical adequacy. In order to be minimally successful, logically-imperfect agents must clarify the epistemic structure of their theories. Otherwise, they will be unable to achieve planning adequacy.[17] To clarify epistemic structure, they will need to determine the epistemic dependence relations for all possible problem-types. This will require reformulating their theories and attendant problem-solving procedures. In this way, reformulating is constitutive of the activity of clarifying epistemic structure.

Recognizing the need to clarify epistemic structure justifies the ubiquity of reformulations in science. Reformulations are not merely practically or instrumentally valuable tools for greater truth or empirical adequacy: they are essential for logically-imperfect scientific agents like ourselves. For such agents, reformulating is a constitutive aspect of scientific inquiry. By reformulating, we gain knowledge of a theory's epistemic structure. Reformulating thereby manifests the gaining of this knowledge, as opposed to merely putting us in a position to gain this knowledge. This contrast parallels a similar distinction that Sosa draws between constitutive vs. auxiliary intellectual virtues. On his view, some intellectual virtues are "knowledge-constitutive," as opposed to merely being auxiliary or practically or instrumentally useful for acquiring knowledge (2015, pp. 41–42).

For convenience, I will refer to the aim of clarifying epistemic structure as aiming

[17]My argument does not entail that logical omniscience is a further constitutive aim of science. Logical omniscience is not required for scientific planning adequacy. One does not need to know all logical truths to succeed at science.

for *epistemic suitability*. This terminology emphasizes that clarifying epistemic structure is necessary for having a suitable problem-solving procedure. It also places this aim within the same grammatical category as empirical adequacy. Just as both truth and empirical adequacy are formal features of theories, epistemic suitability admits a formal characterization. A theory formulation is *epistemically suitable* for solving problems of type P provided that it provides a problem-solving procedure for determining answers to P-problems. Some theory formulations are epistemically suitable for solving particular problem-types, but not others. In order to solve a particular problem using a theory, sometimes it is necessary to reformulate, leading to an alternative theory formulation. In this way, differences in epistemic suitability are analogous to different computer algorithms. One algorithm might be able to solve a problem that another cannot, despite the algorithms being in some sense aspects of "the same theory." To solve the problem using the latter algorithm, we must reformulate it, resulting in a new algorithm.[18]

Some aspects of scientific practice are instrumentally valuable: they are means to some further end. The ends themselves have *final value*: they are valuable in and of themselves, at least viewed from the lense of science.[19] I take it that the constitutive aims of science qualify as final ends of scientific inquiry. For instance, if science constitutively aims at truth, then truth acquires the status of being an end of science. Learning a truth would then have not only instrumental but also final value. In general, being a constitutive aim seems sufficient for being an end in itself (relative to that domain or activity).[20] Therefore, given that epistemic suitability is a constitutive aim, it too is an end of science. By constituting the achievement of this aim, reformulations thereby accrue a form of final value, as opposed to instrumental value alone. Gaining knowledge of epistemic dependence relations is non-instrumentally valuable, in the same way that making a theory more empirically adequate or true is non-instrumentally valuable. I take this to be the most promising argument against instrumentalism about reformulations.

[18]For an illuminating discussion of algorithms, see Goldman (2017, 22ff.). Although seeking testimony from experts *prima facie* counts as a solution procedure (e.g. writing a program that queries Wikipedia for answers), I view this as a subsidiary, practical aspect of scientific methodology. The ability to seek expert testimony is derivative on experts themselves having a non-testimonial solution procedure that articulates the relevant EDRs.

[19]For this distinction between final vs. instrumental value, see Korsgaard (1983).

[20]If science itself is only instrumentally valuable, then the ends of science are some further means to some further end. Settling foundational questions about what—if anything—possesses final value *tout court* lies outside the scope of my discussion here.

As a formal feature of theories, epistemic suitability is not a practical matter. At least, it is no more practical than truth or empirical adequacy. A theory formulation is either epistemically suitable to solve certain kinds of problems, or it is not. It either provides a solution procedure for these problems, or it does not. Nothing in this analysis depends intrinsically on practical features. Likewise, the constitutive aim of epistemic suitability is not a practical aim. At least, it is no more practical than the aims of truth or empirical adequacy. It is a component of the minimal success criteria for science. Insofar as science is an activity, these aims of course make reference to agents. Nevertheless, they do not rely on any special interests or idiosyncratic features of agents beyond their desire for doing science. In Section 5.7, I will contrast these constitutive aims with practical features of scientific inquiry, such as problem-solving speed. First, I consider and rebut objections that threaten to make epistemic suitability a pragmatic aim after all.

5.6 Is Epistemic Suitability a Pragmatic Aim?

Since the aim of epistemic suitability arises for logically-imperfect agents, one might worry that it is inherently pragmatic. This aim appears to arise from a feature of agents, namely their logical imperfection. Yet, the aims of truth or empirical adequacy are no less pragmatic in this regard.[21] As discussed in Section 5.3, the constitutive need for a true or empirically adequate theory also comes from features of most agents, namely that they are neither physically nor observationally omniscient (they have not yet achieved planning adequacy).

Rather than view these aims as arising for agents in particular, contingent epistemic situations, it is better to view them as universal aims of science. The constitutive aims of science apply to *any agent* engaged in science, no matter their epistemic situation. It is just that some agents already trivially satisfy certain constitutive aims. Observationally-omniscient agents trivially satisfy the aim of empirical adequacy. Logically-omniscient agents trivially satisfy the aim of epistemic suitability. Their commitment to science binds them to this aim, while particular features of their cognition conveniently discharge it.

[21]Indeed, one might worry that empirical adequacy *does* fundamentally appeal to agents, since what counts as 'observable' is relative to an epistemic community. Neither truth nor epistemic suitability depend on agents in this way. In response, a constructive empiricist could attempt to define observability in terms of the visual system of a computer.

Human scientists are bound to the same aims, without such cognitive advantages.

One might object that unlike these other aims, epistemic suitability references problem-solving, and problem-solving is an aspect of the *use* of a theory and to that extent pragmatic. Addressing this further worry requires distinguishing between (i) pragmatic (or practical) matters and (ii) (non-practical) epistemic matters. I will consider a few ways of drawing these contrasts. The upshot will be that if "pragmatic" is defined too broadly, then many matters of epistemic interest are pragmatic. Regardless, I take it that we can meaningfully distinguish between non-practical and practical epistemic matters (even if these matters are in some deeper sense 'pragmatic,' a claim that some pragmatists would endorse (Brandom 2011, p. 58)). It is at least in this sense that the aim of epistemic suitability is not pragmatic or practical, but distinctively intellectual (i.e. concerning non-practical, epistemic matters).

van Fraassen's Characterization

In developing constructive empiricism, van Fraassen posited a narrow conception of the epistemic and a broad conception of the pragmatic. According to van Fraassen, epistemic virtues bear on the truth or truth-conduciveness of a theory, concerning "how much belief is involved in theory acceptance" (1980, p. 4). In contrast, any virtue related to people counts as pragmatic, including how we apply theories. Specifically, van Fraassen says that "pragmatic virtues" concern "the use and usefulness of the theory; they provide reasons to prefer the theory independently of questions of truth" (1980, p. 88).[22] Likewise, he characterizes a "pragmatic factor" as "any factor which relates to the speaker or audience" (1980, p. 91).

On this construal of the pragmatic vs. the epistemic, the aim of epistemic suitability does seem to count as pragmatic (but so does much else, as we will see). Epistemic suitability concerns whether or not a theory formulation supports a problem-solving procedure for a given problem. Problem-solving is clearly an aspect of the use or usefulness of a scientific theory. At first glance then, constructive empiricism treats epistemic suitability as a pragmatic virtue.

However, matters are not so simple. In distinguishing the pragmatic from the epis-

[22]Earlier, van Fraasssen claims that "pragmatic virtues do not give us any reason over and above the evidence of the empirical data, for thinking that a theory is true" (1980, p. 4).

temic, van Fraassen is explicitly focused on the context of theory acceptance, where we are choosing between rival theories. This context is importantly different from that which arises when assessing compatible formulations. For in this context, we are not choosing between rivals. Instead, we are assessing whether we can gain knowledge of the world with a given theory formulation. If a particular theory formulation is not epistemically suitable for a given class of problems, then we cannot actually solve these problems using that formulation (we might need a different formulation). This results in a difference in what agents can gain knowledge about. If such differences in knowledge acquisition don't qualify as epistemic, then I don't know what does. It may be that differences in epistemic suitability do not always matter for non-pragmatically choosing between rival theories. Yet, there are other contexts of epistemic assessment besides this one.

Virtue Epistemology

Moreover, defining 'pragmatic' so broadly in all contexts leads to counterintuitive consequences. In particular, it implausibly restricts what counts as non-practically epistemic (i.e. intellectual). Standard assumptions in virtue epistemology provide a particularly dramatic illustration of this problem. In its various forms, virtue epistemology focuses on the intellectual virtues of believers, where these virtues can include abilities, dispositions, competences, or character traits. Here, a 'competence' can be understood as "a disposition to succeed in a given field of aimings" (Sosa 2015, p. 2). Knowledge is then understood as true belief arising from exercising one's intellectual virtue, e.g. an epistemic competence like apt perception. Notice that acquiring knowledge involves an agent to apply an ability. Consequently, van Fraassen's broad characterization of the 'pragmatic' seems to render all such epistemic acts as pragmatic.

Does this make virtue epistemology inherently pragmatic? Perhaps, but either way we can distinguish the intellectual from the practical even within virtue epistemology. We can meaningfully distinguish between virtues that seem constitutive of knowledge vs. virtues that seem merely practically or instrumentally valuable for acquiring knowledge. For instance, Sosa argues that a constitutive feature of some competences is that they manifest knowledge. Knowledge-constitutive competences are purely intellectual, as opposed to practical or auxiliary competences, such as being industrious rather than lazy when it comes to gathering evidence. This contrast underwrites a notion of "*purely* *intellectual* virtues, with no admixture of practical assessment" (Sosa 2015, p. 45).

The thrust of this response does not require adopting virtue epistemology. The point is rather that there are compelling ways to contrast epistemic and pragmatic matters that van Fraassen's 1980s characterization rules out in principle. I take this to indicate that we should not define 'pragmatic' so broadly, at least not when teasing apart practical issues from non-practical epistemic matters (even within a context that itself has pragmatic factors).

Muller's Characterization

Instead, I favor a broader notion of the epistemic and a narrower notion of the pragmatic, such as that proposed by Muller (2005) in his rendition of constructive empiricism. Personally, I would prefer to stop contrasting the epistemic with the pragmatic, and rather just talk about practical vs. non-practical dimensions of the epistemic. But the dominant words are already in use, so there is little I can do to alter settled conventions. Muller defines epistemic aspects of science as those features that are always relevant for deciding whether a proposition of an accepted theory counts as knowledge (2005, p. 63). Epistemic aspects include at least evidence, truth, and empirical adequacy. Unlike van Fraassen's characterization, Muller's entails that many aspects of problem-solving are genuinely epistemic. Figuring out what we need to know or what suffices to know to solve a problem clearly matters for deciding whether a possible solution is a genuine solution. In this sense, epistemic dependence relations are genuinely epistemic. Similarly, differences in epistemic suitability matter for deciding whether to accept a putative solution as knowledge.

In contrast, Muller defines *purely pragmatic* aspects as features that are *never* involved in deciding whether a proposition counts as knowledge. These include aspects of convenience, speed, and efficiency, considered in Section 5.7. More generally, pragmatic features are those that are *not always* relevant for deciding whether a proposition counts as knowledge.[23] On this narrower definition of the pragmatic, only some features of the use or usefulness of a theory are pragmatic. Again, since matters of epistemic suitability are involved in determining whether propositions count as knowledge, Muller's definition

[23]As Muller notes, constructive empiricists—but not scientific realists—view explanation and inference to the best explanation as pragmatic features.

properly classifies epistemic suitability as epistemic rather than pragmatic or practical.

In sum, the applicability of a theory has a distinctively epistemic, non-practical dimension. Problem-solving is not inherently pragmatic. We often use theories to provide reasons for believing or accepting claims about the way the world is. Theories provide these epistemic reasons largely by solving problems. Problem-solving is a central component of the epistemic role of theories. Moreover, the activity of problem-solving seems to be no more pragmatic than the activity of constructing a theory. Indeed, a common way of constructing theories is by unifying and connecting families of problem-solving procedures. If theory construction is not inherently pragmatic, then these constituent problem-solving procedures should not be inherently pragmatic either.

Still, one might worry that epistemic suitability simply tracks what an agent can or cannot do *in-practice*, where such considerations are inherently pragmatic. Given that an agent already has a true theory, can't they already solve *in-principle* any problem they might encounter? If their theory is futuristically true, it already in some sense "contains the answer" to any scientific question they might ask.[24] But truth alone (or empirical adequacy) does not entail that this solution is manifest or available, given your theory formulation. Assuming that epistemic reasons are *reasons for belief or acceptance*, the theory itself does not necessarily provide sufficient epistemic reasons to believe the answer it contains. To acquire these epistemic reasons, it is necessary to formulate the theory such that it is epistemically suitable for solving the given problem. This point is no more pragmatic than needing an (approximately) empirically adequate theory in order to have epistemic reasons in the first place.

To better understand the constitutive, non-pragmatic need for epistemic suitability, consider problem-solving contexts where reformulating is necessary. Perhaps your current theory formulation requires that you perform certain measurements in order to solve a given problem. What happens if you can't perform these measurements? This might be the case if your measurement device breaks, and—as a result of cosmically bad luck—remains broken forever. More prosaically, you simply might be set the problem: solve such-and-such problem without appealing to information that your current theory formulation requires. These are physically possible problem-solving contexts, and one ought

[24]Likewise, an empirically adequate theory is guaranteed to have a model whose empirical substructure is isomorphic to the relevant observable phenomena (van Fraassen 1980, p. 64).

to plan accordingly. Scientific activity involves not just planning for when things go well, but for when things go as poorly as possible. In these circumstances, reformulating your theory is *the only way* to make it applicable. This is not something that you merely have to do 'in-practice:' you have to do this in order to solve the given problem (or to know if it can still be solved at all). If instead you were logically omniscient, then you would already have at your disposal all non-trivially distinct theory formulations. There would be nothing further for you to do. But for logically-imperfect beings, there is a real, in-principle need to reformulate. Reformulations provide contingency plans: they offer alternative epistemic routes to a solution that although 'already encoded' in a true or empirically adequate theory, is not necessarily manifest or accessible.

5.7 A Need for Speed?

Are there any further constitutive aims of science, arising from further constitutive features of science? It is possible that there are. Indeed, my argument for an overarching constitutive aim of scientific planning adequacy—and its subsidiary aims—makes no claim to completeness. In particular, one might wonder whether agents of finite lifespan have a constitutive aim to solve problems as quickly as possible. Although one can pose purely theoretical questions about problem-solving speed or complexity, I will argue that these are separate from practical considerations that might motivate agents to solve problems quickly. Hence, solving problems as quickly as possible (or within other computational constraints) is never a constitutive aim of science, for anyone.

Consider scientific agents with finite computational resources, such as agents of finite lifespan. Is solving problems with minimal resources (or as quickly as possible) a constitutive aim for such agents? As an analogy, consider playing chess within finite time (as most chess is played). Given that a chess player wants to win, they have a constitutive aim to try to checkmate before their time expires. Likewise, given that a human scientist wants to solve a scientific problem, they have a constitutive aim to try to solve it within their lifetime. In both cases, these are constitutive aims *at the agent-level.* As discussed in Section 5.2, both realism and empiricism focus on constitutive aims *at the activity-level*, i.e. the constitutive aims of Chess or Science. Constitutive aims at the agent-level do not entail constitutive aims at the activity-level. A feature or success criterion can be constitutive for an agent without being constitutive of the activity at large.

Considered anthropomorphically, the game of Chess does not care which player wins (or if any player wins at all). Likewise, Science does not care which scientist solves any particular problem, or when. Hence, the constitutive need for a finite agent to try to solve problems as quickly as possible does not entail a constitutive need for speed at the activity-level. As further support for this claim, recall that the ideal constitutive aims of science are futuristic: the minimal criteria for ideal success concern an infinite time frame, with limitless computational resources. Within this futuristic idealization, we treat the activity as having an arbitrarily-long time frame. Even though particular games of chess or scientific problem attempts are finite, the futuristic, ideal aims abstract away from temporal constraints. The constitutive aim of science remains planning adequacy. The computational limits of individual scientists do not affect this constitutive aim, nor do they create an additional constitutive aim.

Although maximizing problem-solving speed or minimizing computational resources are not constitutive aims of science, planning adequacy already encompass related theoretical questions about speed or resource minimization. It *is* constitutive of science to seek knowing how quickly a problem can be solved, along with knowledge of a procedure for solving it that quickly. Questions of speed or resource consumption are themselves theoretical questions, such as *what is the fastest physically possible solution to this problem?* These theoretical questions correspond to problem-solving contexts where the problem is to solve a given sub-problem within a certain period of time. Since scientists ought to prepare for any possible problem-solving context, they ought to investigate the speed at which they can solve problems. The question of how quickly a problem can be solved is yet another scientific question, generating its own problem to solve. These investigations are a purely theoretical aspect of science (and mathematics), with attendant literatures, e.g. on the minimization of proof length or the limits of computer speed. Scientists ought to investigate questions about speed and computational resources, regardless of any practical benefits such investigations might yield.

Lest there remain a whiff of paradox, it is vital to distinguish two different kinds of goals: (i) wanting to know how quickly a problem can be solved vs. (ii) wanting to solve a problem as quickly as possible. The former is a purely theoretical aim (and includes wanting knowledge of how to solve a problem as quickly as possible—e.g. if the world demands it). The latter is a genuinely practical or pragmatic aim. Science itself does not require one to care about how quickly they will actually solve problems (as opposed to how quickly a problem could be solved). Of course, the motivations for asking questions about speed are often practical, borne of a desire to have the fruits of an investigation sooner rather than later. Nonetheless, we can do science as slowly or quickly as we please, while doing science all the same. As Peirce notes, it is constitutive of inquiry to hope that there is an answer to our problem.[25] It is not constitutive (of collective inquiry) to hope that you yourself will find that answer within your lifetime, although you very well might hope this.

One might have thought that desiring speed is non-constitutive because speed is species-relative. What counts as fast for one kind of finite being might be hopelessly slow for another. Yet the empiricist notion of observability is already species-relative, as are the various grades of logical imperfection. Due to species-relativity, the content of the aim of empirical adequacy is necessarily indexical. So it is not this indexical nature which makes a desire for speed into a practical aim, preventing it from being a constitutive aim of science. Instead, a desire for speed is non-constitutive because it arises from the extra-scientific preferences of particular agents, relative to their particular constraints. The only preference that matters for the constitutive aims of science is the desire to do science. Some agents might aim to solve scientific problems quickly or within a reasonable time, but these are practical goals. They typically arise from contingent social structures, such as temporal requirements for career advancement. A finite scientific agent will labor under more practical constraints than an infinite one, but their constitutive aims remain the same. The need for speed is at most a non-constitutive practical aim, at the agentive level.

5.8 Why Talk about Aims at All?

Some may be skeptical that it makes sense to talk about science as having aims. Such worries might stem from a more general skepticism toward 'the aims of activities,' such as

[25]Misak quotes Peirce as saying that "the only assumption upon which [we] can act rationally is the hope of success," which she interprets as meaning that "it is a regulative assumption of inquiry that, for any matter into which we are inquiring, we would find an answer to the question that is pressing in on us" (2013, p. 50).

belief, inquiry, or assertion. Considering the so-called aim(s) of belief, Wedgwood rightly quips that "Beliefs are not little archers armed with little bows and arrows: they do not literally 'aim' at anything" (2002, p. 267). The same can of course be said for science: if it makes sense to talk collectively about science at all, science is not literally aiming at anything. Nonetheless, Wedgwood goes on to give a normative explication of claims about the aim(s) of belief: such claims state correctness conditions for beliefs. Here, I have pursued a similar strategy: constitutive aims of science provide criteria for minimal success in scientific endeavors.[26]

A common objection to aims-talk is that it presupposes a univocal or hegemonic aim, where none is to be found. In his earlier work, Putnam expressed this criticism in a couple places, remarking that "it is hard to believe that there is such a thing as 'the aim of science'—there are many aims of many scientists" (1979 [1971], p. 355).[27] Earlier, he remarked that "The use of such expressions as 'the aim of science'...is already extremely apt to be misleading. For there is no *one* 'aim of science'....Different scientists have different purposes" (1979 [1965], p. 233). More recently, MacFarlane has expressed a similar worry in the context of the aim of assertions. Considering the idea that assertions uniquely aim at the truth, MacFarlane says, "This idea is pretty obscure anyway. Even if truth is *an* internal normative aim of assertion, it is certainly not the only such aim" (2005, p. 227).

In developing my account of constitutive scientific aims, I have not assumed that science has a unique aim. Instead, I have focused on identifying a constitutive aim of science, namely, planning adequacy. I happily allow that science might have other constitutive aims and of course many other non-constitutive aims. As a practical matter, scientists presumably are trying to solve their most pressing problems as quickly as their energy and resources allow.

Part of Putnam's misgivings arise from the fact that individual scientists can themselves have many aims. I take this part of Putnam's worry to lose its force once we distinguish between aims at the agential-level vs. at the activity-level, as discussed in Section 5.2. As I hope to have made clear, constitutive aims of science are largely insulated from the aims of individual scientists. In order for there to be constitutive aims of

[26]More broadly, Sosa argues that the notion of a constitutive aim is useful for understanding action, perception, and knowledge: "We find unity across action, perception, and knowledge. All three are constituted by aimings, by performances with a constitutive aim" (2015, p. 24).

[27]I thank Gordon Belot for suggesting I review this essay by Putnam.

science, it suffices that there are scientific agents aiming to do science.[28]

The primary reason to talk about the aims of science is so simple that it is worth reiterating. In order to determine what it means for a scientific activity to succeed, we must talk in terms of aims, goals, or purposes. These aims or goals define success criteria. As van Fraassen notes, disagreements about "the aim pursued in science" lead to disagreements about what "counts as scientific success" (1995, p. 143). Constitutive aims play a special role in characterizing minimal success criteria for an activity. In this vein, Putnam himself invokes "the aims of inquiry" in order to criticize conventionalist approaches to logic, his point being that we need assurances that such approaches do not interfere with the aim of having "a true description of the world" (1979 [1968], p. 188).

If an activity like science does not have criteria for minimal success, then it loses much of its intellectual interest. If science doesn't aim at anything, then scientists might as well be flailing their arms around. And what intellectual grounds would we have to criticize them? Insofar as we do have grounds to criticize methodologically-wayward scientists, I take there to be a meaningful notion of constitutive aims for science. When we talk about scientists who are failing to do science, we presuppose that there is at least some overarching aim(s) to which such scientists are failing to contribute. Such scientists are not even contributing to the minimal success of science.

Hence, I do not see talk of aims as being any more problematic than talk of correctness conditions, or success conditions more generally. Insofar as particular practices have standards for success, those practices have aims. There are many philosophical concepts that I find deeply mysterious. Aims-talk is not one of them.

5.9 Conclusion

I have argued that one constitutive aim of science is to solve (all possible) problems about the physical world. As a constitutive aim, this specifies a minimal success criterion for science. Scientific realists contend that minimal success requires a true scientific theory. Antirealists contend that truth is not necessary for minimal success. In particular, constructive empiricists argue that empirically adequate solutions to scientific problems

[28]Plausibly, we do not even need this much. We can talk about the constitutive aims of chess even if no one plays chess ever again. Constitutive aims characterize criteria for minimal success in an activity *were there to be agents undertaking said activity.*

would suffice. Since resolving this contentious issue is unnecessary for my larger argument, I have remained neutral on the specific minimal criteria for solving a scientific problem.

Instead, I have argued that—as presently formulated—both realism and antirealism neglect a further constitutive aim of science. Planning adequacy requires more than an empirically adequate or even true theory. In order to solve problems, we require an epistemically suitable formulation of a relevant theory or problem-solving procedure. The aim of epistemic suitability requires that we identify a sufficient knowledge-base for reaching a solution. Ideally, we should identify what we need to know in order to solve scientific problems. Determining what we need to know or what suffices to know to solve problems requires that scientists reformulate their theories and problem-solving procedures. Reformulations are thereby essential for attaining epistemic suitability. For this reason, reformulations are non-instrumentally valuable. They are valuable as scientific ends in themselves, not merely in the service of greater truth or empirical adequacy.

Each constitutive aim has both ideal and non-ideal dimensions. The non-ideal constitutive aims of science concern what science has to achieve for minimal success within its actual history. Non-ideally, we require epistemically suitable formulations only for the problems that we will actually face. Solving problems non-ideally requires only approximate truth or approximate empirical adequacy, where our actual practices settle what counts as 'good enough.' The ideal constitutive aims of science specify what science would have to achieve for minimal success in the limit of infinite time and resources. Ideally, science would be able to solve any physically possible problem about the physical world. This would require complete epistemic suitability for all physical problem-types. Additionally, ideal solutions would be as true or empirically adequate as physically possible, e.g. up to the limits of physically possible measurement precision.

We can distinguish the constitutive aims from additional non-constitutive aims. No doubt, science would go better if we developed faster problem-solving procedures. No doubt, it would be a great success if we arrived at a fundamental language for describing reality. Although these achievements would constitute scientific success, they are arguably not minimal success criteria for science.[29] They can be ideal aims, without being

[29]Fundamentalist scientific realists might argue that seeking a canonical language for reality *is* a constitutive aim of science. I take the arguments in Section 2.5 and Chapter 6 to undermine fundamentalism.

ideal constitutive aims. If constructive empiricism is correct, then truth is also a non-constitutive aim, along with providing explanations of physical phenomena. My defense of epistemic suitability advantageously accounts for the value of reformulating without appealing to explanations or differences in explanatory goodness. It thereby provides an account that realists and antirealists alike can endorse.

Chapter 6:

Making it Manifest: The Intellectual Value of Good Variables

6.1 Hidden Symmetries and Manifest Properties

When discussing the symmetries of models, physicists and chemists sometimes speak of "hidden symmetries." These are symmetries of the model that certain choices of variables obscure. A system possesses a hidden symmetry when its full symmetry group is larger than its "apparent" or "obvious" symmetry group. Paradigmatic examples include the classical and quantum two-body problems (which have a hidden hyperspherical symmetry) and the isotropic harmonic oscillator (which has a hidden special unitary symmetry). A more recent example occurs in the context of $\mathcal{N} = 4$ super Yang–Mills theory, whose tree-level amplitudes possesses a hidden dual superconformal symmetry, along with a larger hidden symmetry known as the Yangian.

By reformulating these models, physicists were able to make these hidden symmetries manifest. The process of making a symmetry manifest distinguishes hidden symmetries from their "obvious" counterparts: non-hidden symmetries were already made manifest in a prior formulation. In some cases, a symmetry is manifest because it is "worn on the sleeves" of a relevant expression. As we will see, this notion of *wearing a property on the sleeves* is a special case of making a property manifest.

The phenomena of manifest symmetries suggests a problem for conceptualism that fundamentalism avoids. *Prima facie*, it is intellectually significant to make a hidden symmetry manifest.[1] It does not seem to be merely a convenient re-expression of a theory or model's known properties. Yet, it is initially not clear how conceptualism can accommodate the intellectual significance of making a symmetry manifest. This is because there

[1] Recall that 'intellectual significance' is a non-practical dimension of epistemic significance.

is typically a translation procedure between variables that obscure a symmetry and variables that make this symmetry manifest. Hence, it initially seems that both sets of variables must express the same set of epistemic dependence relations. If this were so, then conceptualism would fail to save the intuition that something of intellectual importance can occur when scientists make a symmetry—or other property—manifest.

In contrast, fundamentalism suggests a simple account of the intellectual significance of making properties manifest. Expressive means that make more fundamental properties manifest carve nature more closely at the joints. Insofar as a symmetry qualifies as fundamental, making it manifest would likewise count as being intellectually significant. Indeed, symmetries are connected with physical invariants, and invariants are typically taken to be physically fundamental. If conceptualism cannot provide a satisfying account of making symmetries manifest, it would seem as though fundamentalism has the upper hand in this context. This chapter provides a conceptualist account of the significance of making properties manifest, including symmetries. In keeping with the methodological desiderata of Chapter 2, I will not appeal to ontologically-primitive differences in joint-carving or fundamentality. Such differences might obtain, but I will remain agnostic as to whether they do. Instead, I will locate a source of non-practical, epistemic value in making properties manifest.

Section 6.2 begins with a general account of what it means for a fact to be manifest rather than hidden. I then consider the ubiquitous phenomena of expressions that wear a property "on the sleeves." Section 6.2.1 analyzes this as a special case of making a property manifest. I illustrate my account with simple examples from math, physics, and logic. Next, Section 6.3 applies my account to a simple example from language translation: some languages make the meaning of a word more manifest than others. Section 6.5 considers the more complicated but still prosaic context of coordinate transformations.

In all of these cases, conceptualism threatens to either collapse into instrumentalism or risk expanding into fundamentalism. For instance, if no coordinate choice carves the system more closely at its joints, then how can we intellectually privilege one set of coordinates over another? For many coordinate transformations, there seems to be nothing but convenience to decide between them. Section 6.4 responds to these worries by clarifying the non-practical epistemic value of making properties manifest. Making a property manifest is valuable whenever it rules out epistemically possible solutions to a given problem. This ruling out of possibilities has epistemic value independently of any practical value. Since I do not appeal to primitive differences in fundamentality, my account shows that fundamentalism is not needed even in this context. My argument complements Woodward's (2016, p. 1056) argument that appeals to joint-carving do not help us resolve philosophical problems about good variable choice.

Nevertheless, a fundamentalist might object that my account fails to preserve ordinary judgments regarding relative fundamentality. Physicists and mathematicians commonly view some variable choices as being more fundamental or deeper than others. Fundamentalism seems well-suited to vindicate these ordinary judgments of fundamentality. In contrast, conceptualism faces the burden of accounting for them without appealing to substantial metaphysical commitments. To meet this burden, Section 6.6 proposes an expressivist account of fundamentality. To judge that a formulation X is more fundamental than a formulation Y is to express a mental state of *being for privileging X over Y*. Using the example of gauge choices in quantum field theory, Section 6.7 develops a separate argument against fundamentalism. Making one fundamental property manifest often comes at the cost of obscuring others. This provides some reason to be pessimistic that physics will ever arrive at a fundamental language that avoids these trade-offs.

I end by considering examples that have motivated the entire enterprise: hidden symmetries. Section 6.8 illustrates my framework in the context of the hidden hyperspherical symmetry of the nonrelativistic hydrogen atom. In many formulations of the hydrogen atom, this symmetry is hidden while hydrogen's spherical symmetry is manifest. By moving to momentum space, we can make this hidden $SO(4)$ symmetry manifest. Finally, Section 6.9 considers hidden symmetries in the context of $\mathcal{N}=4$ super Yang–Mills theory. I describe the chain of variable choices that allow us to make a hidden dual superconformal $SU(2,2|4)$ symmetry manifest. At each step in this long chain of variable changes, we acquire intellectually significant benefits.

6.2 Manifest vs. Hidden Facts

To account for the wide variety of cases that interest me, I propose the following account of manifest facts: a fact is *manifest* at a given stage in a problem-solving plan provided that an agent who implements that plan ought to infer that fact. More precisely:

> **Manifest fact**: a fact F is manifest in epistemic circumstance C provided that an agent in state C ought to infer that the fact F obtains.

On this characterization, solutions are always manifest at the end of a successful problem-solving plan: an agent that implements the plan ought to infer the solution. I take this feature to be a conceptual requirement of any definition of 'manifest fact.' It is constitutive of a successful problem-solving plan that it makes the solution manifest. Otherwise, the plan has not reached its aim and to that extent remains unsuccessful.

By 'agents,' I mean to include both sapient and non-sapient problem-solvers, such as algorithms implemented by a computer program. Sapient agents have a further capacity for grasping a problem-solving plan, thereby understanding it in a psychological sense. Sapient agents can not only implement a plan but also understand it.

My characterization of manifest facts treats it as a normative aspect of problem-solving plans. Whether or not a fact is manifest depends on what we epistemically ought to infer. Some may be wary of normativity, but there is nothing to fear, even for a hard-nosed empiricist or naturalist like myself. Gibbard (2012) provides an ontologically non-mysterious account of what constitutes these ought-claims. They simply amount to plans for action or belief. In particular, "epistemic ought beliefs amount to plans for degrees of credence" (2012, p. 178). To simplify the discussion, I will typically talk in terms of full-belief, although it is straightforward to generalize the account to degrees of credence. Degrees of credence accommodate problem-solving plans that involve inductive rather than deductive reasoning.

We can likewise characterize what it means for a fact not to be manifest, i.e. to be non-manifest. We simply negate the characterization of a manifest fact:

> **Non-manifest fact**: a fact F is not manifest in epistemic circumstance C provided that it is not the case that an agent in state C ought to infer F.

For instance, solutions are not manifest at the beginning of problem-solving (otherwise, one would not need to engage in problem-solving). It is not the case that one ought to infer the solution to a problem before carrying out an adequate problem-solving plan.

That a fact is not manifest does not necessarily entail that it is hidden or obscured. It may sometimes be permissible for an agent to infer a fact that is not manifest. To characterize what it means for a fact to be hidden, I propose the following logically stronger definition:

> **Hidden fact**: a fact F is hidden in epistemic circumstance C provided that it is *impermissible* for an agent in state C to infer that F obtains.

Equivalently, a fact is hidden provided that an agent ought not infer it.

Epistemic-ought claims play an important role in my account of manifest, non-manifest, and hidden facts. But what does it mean to say that an agent in a particular circumstance *ought to infer* a given fact? We can gloss this as follows: if an agent ought to infer F, but they fail to infer F, then their inferential omission warrants disapproval. This disapproval is of a specifically epistemic variety: it is disapproval on epistemic grounds. In the cases I consider, it involves disapproval of the agent's subsequent epistemic state.[2]

If an agent fails to infer the correct answer, then they either (i) infer an incorrect answer, (ii) fail to realize that they know how to solve the problem (e.g. by falsely believing that they do not have enough information), or (iii) simply do not know how to solve the problem. The first two cases involve a kind of epistemic mistake: the agent believes something false (either the wrong answer or an erroneous belief about what is possible). In the third case, the agent displays an epistemic deficiency: they are unable to implement an appropriate problem-solving plan. Of course, this third case warrants disapproval only if the agent ought to know better, i.e. ought to be able to implement the plan. In the cases I consider, I will assume that the agent either knows or ought to know how to implement such a plan. A computer program can malfunction in all three of these different ways. It might halt at the wrong answer, fail to halt when it should, or simply stop working entirely (and not because it has been turned off!).

A simple example from graph theory illustrates the various components of my account. Given a graph (i.e. a collection of edges and vertices), one general question is whether the graph has the property of planarity. Planar graphs admit a representation such that no edges cross in the plane. For any given planar graph, most of its representations hide the fact that it is planar. These representations hide the planarity of the graph by representing two or more edges as crossing. In contrast, other representations demonstrate that a planar embedding is possible: they make manifest the planarity of the graph.[3]

---

[2]In general, we might also epistemically disapprove of an agent's epistemic process. For instance, we disapprove of an agent who gets the right answer but for the wrong reasons, such as by accidental cancellation of two compensating mistakes.

[3]Elsewhere in pure mathematics, number theory provides numerous cases where reformulating a problem makes otherwise hidden patterns manifest (Ash and Gross 2008).

If a student of graph theory is shown a planar representation of a graph, the student ought to infer that the graph is planar. If they do not make this inference—drawing some other inference instead—then their inference warrants epistemic disapproval. For they have either i) inferred that the graph is not planar, ii) inferred that there is not enough information to solve the problem, or iii) realized that they don't know how to solve the problem. In the first two cases, they make an epistemic mistake. In the third case, they display a deficiency that they ought not have (given their background training in graph theory). They show that they lack sufficient understanding of graph theory, whereas they ought to have this understanding.

Of course, if it is not the case that an agent ought to have this background knowledge, then they make neither an epistemic mistake nor display an inexcusable epistemic deficiency. If you show a kindergartner a planar representation of a graph and ask them whether the graph is planar, they can permissibly reply that they have no idea what you are talking about. Although the kindergartner has an epistemic deficiency, they are excused from disapproval. It is not the case that they ought to understand graph theory. Likewise, if someone simply loses interest in solving a problem and walks away, we cannot epistemically disapprove of them for this. We might still, nonetheless, disapprove of their values and goals.

The definitions of manifest, non-manifest, and hidden facts reference an epistemic circumstance C. This circumstance encompasses both i) the background knowledge and capacities that the agent has and also ii) what information they are being presented with in a given problem-solving context. Sometimes, it will be convenient to isolate the latter information, calling it *the problem-specific epistemic circumstance* P. In the case above, both the graph theory student and the kindergartner are presented with the same problem-specific circumstance P, i.e. the same representation of the graph. But overall they are in different epistemic circumstances based on their different background knowledge. The planarity of the graph is manifest for the student of graph theory but not for the kindergartner.

The phenomena of perfect (or absolute) pitch helps illustrate why it is necessary to index what we ought to infer to our background knowledge and capacities. Consider two musicians presented with the same sound, such as a musical note sustained on a violin.[4]

[4]In virtue of being musicians, these agents understand the naming conventions linking frequencies to pitches.

The first musician has perfect pitch. In virtue of this, they ought to infer the pitch class of the note played, e.g. that it is a B-flat. To do this, they do not need any measuring device or even a reference pitch. The second musician does not have perfect pitch. Hence, it is not the case that they ought to infer that the note is a B-flat, just from hearing it. It is epistemically permissible for them not to know the pitch. In order for the pitch to become manifest to the second agent, they need a measuring device, such as a tuning fork, a digital tuner, or testimony. Using a digital tuner alters their epistemic circumstance, such that the pitch of the note becomes manifest. As described below, formulations that "wear a property on the sleeves" are analogous to having perfect pitch. They make it the case that one ought to infer the property without needing intermediary expressions, analogous to how someone with perfect pitch does not need an intermediary measuring device.[5]

The capacity of logical omniscience provides another illustration of how what an agent ought to infer can depend on their capacities. Logically omniscient agents ought to infer any logical consequence of a sentence or group of sentences. For them, all logical consequences are manifest. Clearly, this is not the case for us, in virtue of our lack of logical omniscience. As in Chapter 5, I am interested in agents that are not logically omniscient. Most of the epistemic differences that interest me here do not arise for logically omniscient agents. Unlike humans, such agents would have no reason to reformulate in many of the cases described below.

6.2.1 Simple examples, on the sleeves

My account of manifest facts leads straightforwardly to an account of what it means for an expression to wear a property "on its sleeves." I propose to understand this as follows:

> **To wear on the sleeves ('sleeve properties')**: a representation or expression E wears a property P on its sleeves provided there is a problem-solving plan that both
> (i) makes P manifest and
> (ii) does so solely on the basis of manifest facts about E.

Unpacked, this definition comes to the following: applying an appropriate plan to the

[5] My account of manifest properties sheds light on Wittgenstein's discussion of what he calls "aspect-blindness" in Fragment xi of the *Philosophy of Psychology*. Wittgenstein is concerned with humans that lack "the ability to see something *as something*." He asks whether this "defect" would be comparable "to not having absolute pitch" and later answers that "aspect-blindness will be *akin* to the lack of a 'musical ear'" (2009 [1949], 224–225, §257, 260). In my terminology, aspect-blindness occurs when the changing of aspects—such as the Gestalt shift of the Necker cube—is not manifest to an agent.

expression E makes the property P manifest. Importantly, this plan must rely solely on properties of E that are already manifest (before implementing the plan). Collectively, the expression and the plan generate an epistemic circumstance in which the property is manifest. Typically, these plans are built around a central epistemic dependence relation, which we exploit to make the property P manifest. In this section, I illustrate my account using some simple examples from logic and physics.

Sleeve Properties in Truth-Functional Logic

Sentential logic provides a wellspring of examples of "sleeve properties."[6] Among different but truth-functionally equivalent sentences, often one wears a property on the sleeves that another obscures. Much of the interest in certain kinds of normal forms for truth-functional sentences comes from making certain properties manifest.[7]

The completed truth table of a sentence wears many of the sentence's truth-functional properties on its sleeves. These include whether the sentence is a tautology, a contradiction, or contingent (i.e. true under some but not all truth-value assignments). For instance, to determine if a sentence is a tautology, it suffices to check whether it is true under every possible truth-value assignment to its atomic sentence letters. This epistemic dependence relation supplies an appropriate plan for determining whether a sentence is a tautology. The completed truth table makes manifest the sentence's truth-values, e.g. by collecting them under the sentence's main connective. In other words, given the completed truth table, one ought to infer the sentence's truth-value for every truth-value assignment to its atomic sentence letters. Then, by applying the preceding EDR for a tautology, the truth table makes manifest whether the sentence is a tautology. The truth table thereby wears this property on the sleeves, namely, the property of being a truth table of a tautology. Provided that one sees the truth table and applies this EDR, they ought to infer that the sentence is a tautology.

Some sentences wear their tautological status on their sleeves all by themselves, no truth table needed! As a simple example, consider the sentence $(p \vee \neg p) \wedge (\neg r \vee q \vee s \vee \neg q)$, which is in conjunctive normal form (CNF). To see that this sentence is a tautology, we

[6]Using an ellision introduced below, we could elide "sleeve properties" to the simpler "manifest properties," where this notion would now encompass properties manifest to the 0th or 1st degree.

[7]The following discussion draws heavily upon Goldfarb (2003, pp. 67, 73–74), which first exposed me to the idea of an expression wearing a property on its sleeves.

can rely solely on features of it that are already manifest. For instance, it is manifest that the sentence consists of two conjuncts, each of which is a disjunction of negated and unnegated sentence letters. Anyone who understands the sentence ought to infer these surface-level properties; they are trivially manifest—what we might call 'manifest to the 0th degree.' Moreover, any agent who knows the following EDR also ought to infer that the sentence is a tautology: to determine if a conjunction of disjunctions is a tautology, it suffices to check whether each conjunct contains a sentence letter that occurs both negated and unnegated. In the first conjunct, it is manifest that p occurs negated and unnegated, whereas q occurs negated and unnegated in the second conjunct. Hence, each conjunct is a tautology, so the sentence itself is a tautology. Combined, the sentence and this EDR make manifest that the sentence is a tautology (we might say that this property is 'manifest to the 1st degree'). In general, any sentence in conjunctive normal form wears the property of being a tautology (or not) on its sleeves. We simply use the following epistemic dependence relation: to determine whether a sentence in CNF is a tautology, it suffices to check whether each conjunct contains a sentence letter and its negation.

A sentence is in disjunctive normal form (DNF) provided that it is a disjunction of conjunctions of sentence letters or their negations, such as the following sentence: $(p \wedge q) \vee (\neg s \wedge r) \vee (\neg p \wedge p)$. DNF makes manifest whether a sentence is *satisfiable*, i.e. is true on some truth-value assignment. This follows from logical properties of disjunctions and conjunctions. A disjunction is satisfiable if and only if it has a satisfiable disjunct. In DNF, each disjunct is a conjunction. Hence, we note further that a conjunction is satisfiable if and only if it is not truth-functionally equivalent to a contradiction, such as "$p \wedge \neg p$." Collectively, these two facts yield the following epistemic dependence relation: to determine whether a sentence in DNF is satisfiable, it suffices to check whether at least one disjunct does not contain a sentence letter and its negation. Similarly, to determine whether a sentence in DNF is unsatisfiable (i.e. false on every truth-value assignment), it suffices to check whether every disjunct contains a sentence letter and its negation (in which case every disjunct is a contradiction). Provided we implement these two EDRs, a sentence in DNF wears its satisfiability or unsatisfiability on its sleeves. More generally, disjunctive normal form makes manifest the truth-value assignments on which the sentence is true (it wears these assignments on its sleeves).

Manifest Lorentz Covariance

Physicists commonly refer to some expressions as being "manifestly Lorentz covariant." For instance, the following equation is manifestly Lorentz covariant: $a_\rho a_\nu b^{\rho\mu} = B^\mu_\nu$. This simply means, I will argue, that this expression wears the property of Lorentz covariance on its sleeves. In conjunction with an appropriate EDR, one ought to infer that this expression is Lorentz covariant, *solely on the basis of properties that are already manifest.*

A suitable plan for checking whether an expression is Lorentz covariant comes from the following fact: an equation in tensor form is Lorentz covariant provided that i) non-repeated upper and lower indices on either side match and ii) repeated indices appear once lower and once upper on the same side of the equation. This fact yields the following EDR: to check whether a tensor equation is Lorentz covariant, it suffices to check whether these two conditions are met. Notice that these conditions rely on properties of the equation that are already manifest, namely the occurrence and placement of indices. Hence, an agent who understands this EDR and sees the expression ought to infer that the equation is Lorentz covariant. Likewise for any other expression that satisfies these conditions. Such expressions wear Lorentz covariance on their sleeves. (For a non-conscious agent, we can replace talk of 'seeing' and 'understanding,' with notions of being given the expression as *input* and *implementing* this problem-solving plan.)

My account also illuminates what it means to say that an expression is manifestly Lorentz invariant. When physicists say this, they simply mean that the expression wears Lorentz invariance on its sleeves. An example is the expression $F_{\mu\nu}F^{\mu\nu}$. Here, each lower index is paired with a matching upper index, and there are no free indices. These manifest facts suffice for inferring that the expression transforms as a scalar under Lorentz transformations. Hence, in conjunction with this problem-solving plan, the expression $F_{\mu\nu}F^{\mu\nu}$ wears its Lorentz invariance on its sleeves.

In contrast, some expressions are Lorentz invariant, but this property is not worn on the sleeves (it is hidden). One can prove that such expressions transform as a scalar, but it is not the case that one ought to infer this *solely on the basis of properties that are already manifest.* Instead, one must rely on properties that become manifest only after starting the proof. A well-known example is the Lorentz invariant measure $\int \frac{d^3k}{(2\pi)^3 2w_k}$, where $w_k = +\sqrt{|k|^2 + m^2}$. One can prove that this measure is invariant under proper orthochronous Lorentz transformations. By the end of this proof, its Lorentz invariance is manifest. But the expression itself does not wear this property on its sleeves, in the way that "$F_{\mu\nu}F^{\mu\nu}$" does. At least, I do not know of any appropriate EDR that makes this property manifest solely on the basis of properties of "$\int \frac{d^3k}{(2\pi)^3 2w_k}$" that are already manifest.

These examples illustrate a general epistemic difference between expressions that wear a property on the sleeves vs. those that do not (but that still possess the property). In both cases, to make the property manifest, we must engage in problem-solving. We must apply an epistemic dependence relation(s) that forms the basis of a problem-solving plan. When the property is worn on the sleeves, we do not need to consider any intermediary expressions. The expression itself contains sufficient information for determining whether the property obtains. In contrast, when the property is not worn on the sleeves, we must construct intermediary expressions, such as a truth table. It is from these intermediaries that the property ultimately becomes manifest (i.e. at the end of problem-solving). Section 6.4 analyzes this kind of epistemic difference in terms of a difference in the ruling out of epistemic possibilities. A formulation that makes a property manifest rules out possibilities that the non-manifest formulation does not. This kind of epistemic difference contributes to the non-practical epistemic value of making properties manifest, i.e. to its intellectual significance.

From these considerations, we begin to see how one could construct a gradated notion of manifest properties. Clearly, there is a sense in which properties that are worn on the sleeves are *more manifest* than those that are not. Moreover, we have seen that some properties are trivially manifest, and thereby trivially worn on the sleeves. Above, I referred to these as being "manifest to the 0th degree." Properties that are non-trivially worn on the sleeves are "manifest to the 1st degree": we make sleeve properties manifest by relying on an EDR that exploits only properties that are 0th-degree manifest. Often, when scientists and mathematicians talk about "manifest properties," they really mean properties that are non-trivially worn on the sleeves. Such properties are not immediately manifest, but they become manifest once we apply an EDR that relies solely on already manifest properties. Section 6.4.3 develops a more general proposal for a gradated account of manifest properties.

6.3 Manifest Meanings

As discussed briefly in Section 2.6, languages can differ in how manifest they make the meaning of a word. At first glance, the German word "die Speisekarte" is completely synonymous with the English word "the menu." Both mean what we can denote at the level of thought by 'MENU.' Yet, due to the sub-word structure of "die Speisekarte," German makes the meaning of this word more manifest than English. On my account, this means that there are problem-solving contexts where a German speaker ought to infer the meaning of "die Speisekarte," whereas an English speaker in the same (non-linguistic) epistemic circumstance ought not infer the meaning of "menu."

Consider two agents, Gertrude and Ender, who are native speakers of German and English, respectively. Gertrude has forgotten the meaning of "die Speisekarte" while Ender has forgotten the meaning of "menu." Thanks to the semantic substructure of "die Speisekarte," Gertrude is in an epistemically superior position. In German, "die Speise" means DISH or FOOD, while "die Karte" means CARD or CHART. Hence, Gertrude ought to increase her credence that "die Speisekarte" means a card or chart that displays dishes or food, i.e. that it means MENU. In contrast, Ender is not permitted to make a similar inference. Knowing the meanings of "dish" and "card" is of no use here, since the English word "menu" does not have an analogous substructure. On the basis of what he can remember, Ender ought not increase his credence that "menu" means MENU. Hence, German supports a problem-solving plan that English does not. In virtue of this plan, German makes manifest the meaning of "die Speisekarte," whereas English does not make manifest the meaning of "menu."[8]

In order for Ender to carry out Gertrude's problem-solving plan, Ender would effectively need to 'change variables' by translating into German. For instance, Ender would need to know that the English word "menu" is synonymous with the German "die Speisekarte," and that "die Speise" means DISH/FOOD while "die Karte" means CARD/CHART. Given this additional information, Ender ought to increase his credence that "menu" means MENU. But notice how Ender requires knowledge of a translation procedure, whereas Gertrude does not. This provides another way of seeing that German,

[8]Similar examples of this kind include the German for 'attractions'—die Sehenswürdigkeiten (things worthy of seeing)—and for 'deranged'—geistesgestört (distortion of the mind or spirit). The German for 'attractions' makes manifest that these are things that are typically worth seeing.

but not English, makes the meaning of this word manifest.

Of course, there are other ways to make the meaning of a word manifest. For anyone with sufficient background knowledge, a dictionary makes manifest the meaning of unknown words. Ender could look up the meaning of “menu” in an English dictionary. Its meaning would be made manifest by a definition such as this: “a list from which to request food dishes at a restaurant or social event.” Provided that Ender knows the meanings of enough of these words, he ought to increase his credence that “menu” means MENU. Gertrude could likewise follow this alternative problem-solving plan, consulting a German dictionary for the meaning of “die Speisekarte.”

Perhaps one might worry at this point: is there really any philosophically interesting difference between Gertrude inferring the meaning of “die Speisekarte” from the meanings of its sub-words vs. inferring its meaning from a dictionary? Indeed, in both cases, Gertrude infers the meaning of “die Speisekarte” on the basis of knowing other words. As we have seen above in the context of sleeve properties, there is at least one important difference. Through the former problem-solving plan, the word “die Speisekarte” makes its meaning manifest, solely using features of it that are already manifest (namely, its sub-word structure). A German does not need a German dictionary for this. Whereas in the latter problem-solving plan, a dictionary does the work (indeed, a good dictionary does this work in any language, for any word—at least for speakers with sufficient knowledge of the language). Relative to the problem-solving plan that relies on a dictionary, there is no epistemic difference between Gertrude and Ender. Relative to the problem-solving plan that involves a decomposition into sub-words, an epistemic difference arises.

This example illustrates that it is not the expressive means on its own that makes a property manifest. Rather, it is the expressive means in conjunction with a problem-solving plan. Whether a property is made manifest depends on how one plans to use an expressive means. Ender could make the meaning of “menu” manifest if he chooses to use an English dictionary. But Gertrude does not need a dictionary, provided that she plans to infer the word’s meaning from known sub-words. This example is particularly striking because it shows how the problem-solving plans that are available can depend on the choice of expressive means, e.g. language or notation. Ender is not even able to carry out the sub-word decomposition plan that Gertrude follows. If Ender tries to apply this EDR, it takes him nowhere, for “menu” does not decompose into English sub-words.

6.4 The Value of Making it Manifest

Having expounded my account of what it means to make a property manifest, I return now to this chapter's central question: what is the value of making properties manifest? More precisely, what is required for it to be valuable? This question is a special case of Chapter 2's investigation into the value of compatible reformulations. As before, at least three dimensions of value suggest themselves: instrumental/practical, metaphysical, and non-practical epistemic (what I am calling 'intellectual' value). After presenting instrumentalist and fundamentalist accounts of the value of manifest properties, I propose a conceptualist middle ground.

Both instrumentalism and fundamentalism give straightforward criteria for when it is valuable to make a property manifest. According to instrumentalism, making a property manifest is valuable whenever it contributes to the achievement of other scientific aims. For instance, provided that making a property manifest makes problem-solving more convenient or efficient, it is valuable by the lights of instrumentalism. The instrumentalist denies that making a property manifest ever constitutes *on its own* the achievement of a scientific aim. The most austere form of instrumentalism—conventionalism—contends that making properties manifest is *merely convenient*.

In contrast, fundamentalism contends that making a property manifest can constitute the realization of an aim of science, namely the aim of describing reality in ever more fundamental terms. On this view, a variable choice is valuable at least when it leads to a more fundamental or joint-carving description of a given phenomenon. Consequently, making a property manifest is valuable whenever doing so constitutes a more fundamental description of the phenomena. *Ceteris paribus*, a variable choice that makes a more fundamental property manifest qualifies as more valuable than a choice that obscures such a property.

To fare at least as well as these accounts, conceptualism must provide clear criteria for when it is valuable to make a property manifest. Such criteria must underwrite an evaluative asymmetry between those variable choices that make a property manifest vs. those that do not. The former are better or more valuable than the latter, other things equal. Fortunately, an empiricist-friendly, epistemic criterion lies ready at hand, which I articulate in Section 6.4.1.

As a warm-up, consider the simple case where our epistemic end is to determine whether an expression or system possesses a particular property. By making that property manifest, we achieve our epistemic end. In this context, making a property manifest constitutes the achievement of our goal. Insofar as achieving this goal is epistemically valuable, so is making the property manifest. Variable choices that fail to make the property manifest fall short of this goal, and are to that extent less valuable. Such variable choices do not preclude us from obtaining this knowledge, but they do not suffice for it.

To see this, recall the graph theory example from Section 6.2. By making planarity manifest, we already achieve our aim of determining whether the graph is planar. In virtue of this property being manifest, we ought to infer planarity. In contrast, a non-planar representation of the graph does not suffice for achieving our aim. Doing so requires a further epistemic transformation, such as constructing a planar representation from the non-planar one.

My conceptualist account has important differences with both instrumentalism and fundamentalism. In contrast with instrumentalism, making a property manifest is not *merely* an instrument for achieving scientific aims. Instead, it can constitute the achievement of epistemic aims, such as knowing whether or not a system has a particular property.

Additionally, the kind of epistemic value that conceptualism identifies is logically independent from what is all-things-considered most practically valuable. In the context of determining the meaning of 'menu,' it will typically be more convenient to use an English dictionary than to translate 'menu' into German, learn the meanings of some German subwords, and then translate back. Nevertheless, there remains a sense in which German is epistemically better suited to solve this problem. Similarly, we can imagine contexts where someone with perfect pitch would prefer to use a measurement device to determine the pitch of a sound. Perhaps the sound is extremely loud, and they desire to protect their ears by measuring the sound while in a different room. Hence, the conceptualist criterion for significance has nothing intrinsically to do with speed, convenience, or other practical dimensions of value.

The 20th century Russian physicist Vladimir Fock's commitment to Marxism provides an illuminating historical example. Motivated by dialectical materialism, Fock developed harmonic coordinates as a preferred coordinate system for expressing equations in general relativity (Graham 2000, p. 34). One can imagine it being prudent—in certain political contexts—to prefer Fock's formulation regardless of its epistemic benefits. Vice versa, one might sometimes prefer to use harmonic coordinates for their intellectual advantages, even while denouncing Marxism in all its forms.[9]

In contrast with fundamentalism, conceptualism contends that the value of making a property manifest does not depend on that property being relatively fundamental. In the simple case of checking whether a system has a property, all that matters is that we have a prior epistemic aim of determining whether the system has this property. The conceptualist account applies to any kind of property of interest, regardless of whether such properties are relatively fundamental. The same sorts of epistemic differences can arise for properties that are completely non-fundamental or 'gruified.'

This flexibility presents one of the chief advantages of conceptualism over fundamentalism. At least part of the epistemic value of making properties manifest floats free from the relative fundamentality of those properties. In many contexts, none of the properties that we make manifest seems to be most metaphysically fundamental. Section 6.5 provides a simple example stemming from the choice of Cartesian vs. polar coordinates. It is implausible that one set of coordinates counts as 'metaphysically more fundamental' than another. After all, coordinates are ways of representing states of affairs, rather than properties of those states of affairs. Nonetheless, the kinds of epistemic differences that arise in these cases completely parallel the differences that arise in cases where fundamentalism might get traction, such as the case of hidden symmetries or gauge choices. Yet as Section 6.7 shows in the context of gauge choices in quantum field theory, making one (fundamental) property manifest often comes at the expense of obscuring others. To assess the relative value of these gauge choices, the fundamentalist requires some way of comparing these trade-offs. Conceptualism does not require this kind of accounting in order to make sense of why it can be valuable to make different properties manifest in different contexts.

6.4.1 Ruling out epistemically possible solutions

Of course, in many cases our task is more complicated than simply checking whether or not an expression has a property. Section 6.5 illustrates one such context, where the task

[9] Thanks to Gordon Belot for suggesting this example.

is to determine the equation of a line. For a horizontal line, I argue that it is intellectually valuable to make the vertical degree of freedom manifest, although this is not the same as determining the equation of the line. Hence, we need a more general criterion for the intellectual value of making properties manifest. For ease of discussion, I introduce some terminology: let's call a formulation or choice of variables that makes a property (more) manifest a "(more) manifest formulation." Conversely, let's call a formulation that fails to make a property manifest a "non-manifest formulation" (or at least a "less manifest" one).

In general, a manifest formulation has the following epistemic advantage: *it rules out epistemically possible solutions that a non-manifest formulation does not.* To see this, it helps to reconsider some prior examples. A planar representation of a planar graph rules out the epistemic possibility that the graph is not planar. In contrast, when presented with a non-planar representation, it remains epistemically possible that the graph is not planar. Consider next a person with perfect pitch. In virtue of pitch being manifest to them, they immediately rule out epistemic possibilities that an ordinary person can rule out only via a measurement device. Clearly, there is an epistemic advantage to having perfect pitch, even for someone who chooses not to use it for practical reasons. Similarly, when we present an expression in manifestly covariant form, we rule out the epistemic possibility that the expression is not Lorentz covariant. A non-covariant form leaves open this epistemic possibility, in the sense that it is not the case that we ought to infer that the expression is Lorentz covariant.

Of course, what matters isn't simply the *number* of epistemic possibilities that are ruled out. What matters is ruling out epistemically possible *solutions*, rather than epistemic possibilities *tout court*. Relative to the aim of solving a particular problem, there is little-to-no value in ruling out epistemic possibilities that have nothing to do with that problem. If I am trying to determine whether a sentence is a tautology, I do not advance by noting that I am wearing pink socks (despite the fact that this observation rules out many epistemic possibilities). Indeed, as the examples in Sections 6.5 and 6.7 demonstrate (concerning coordinate choices and gauge choices, respectively), a formulation that is non-manifest with respect to one property can be manifest with respect to another. Hence, if we were to naïvely count the epistemic possibilities that such formulations rule out *tout court*, we would miss key epistemic differences between them.

I am proposing that the value of making a property manifest derives from ruling out possibilities that we could rationally entertain in the course of problem-solving.[10] Unsurprisingly, the class of possibilities that matters changes across different kinds of problems. As Section 6.4.3 discusses, a problem's epistemically possible solutions constitute a 'search space' for that problem. Making a property manifest is epistemically valuable when it constrains this search space.

As noted above, conceptualism does not deny the instrumental or practical value of making properties manifest. Indeed, this practical value often stems from the conceptualist criterion I have just proposed: by ruling out more epistemically possible solutions, manifest formulations are often more convenient for problem-solving. By eliminating these possibilities, it typically becomes easier or faster to solve a problem. Of course, other practical considerations can intervene, such as the pedagogical costs of learning a manifest formulation. Hence, as I have been arguing, these are genuinely independent dimensions of value. Although greater convenience is often a symptom of intellectual significance, it is not a criterion.

Section 6.9 provides a striking illustration of this moral, involving a kind of case that arises frequently with symmetries. By reformulating such that a symmetry is put on the sleeves, we gain the ability to construct increasingly complicated expressions that manifestly respect this symmetry. In this case, we gain the ability to build more complex invariants out of starting points that are manifestly invariant under the symmetry. Unsurprisingly, this ability is incredibly convenient in many contexts. It is so convenient that it is easy to lose sight of its underlying intellectual significance, which obtains independently of these practical benefits. By making the symmetry into a sleeve property, one ought to infer that a given expression has that symmetry. Whereas otherwise, it would be a live possibility that the expression is not invariant. Due to this epistemic possibility, it would be necessary to check—via calculation—that the expression has the symmetry in question. Section 6.9's example, involving supersymmetry, also makes salient the fact that making a symmetry into a sleeve property can take *a lot* of work. In many contexts, it would not be practically worth doing this work, unless one was faced with multiple problems that could practically benefit from it.

[10]Such rationally-entertainable possibilities just are the epistemically possible solutions. If we follow Gibbard (1990) in giving an expressivist treatment of rationality, then judgments about possible solutions are inherently normative.

6.4.2 Less surprising, more intelligible

By ruling out epistemic possibilities, a more manifest formulation makes the phenomena of interest less surprising or mysterious. If we start with a more restricted space of possible solutions, the fact that the solution has a given property is typically less surprising than it otherwise would be. (If we were to apply a principle of indifference, we would begin with different priors concerning the property of interest, depending on whether we start within a more manifest formulation vs. a less manifest one.) I take this decrease in surprise to be a sufficient condition for greater intelligibility. More manifest formulations often make the solution to a problem more intelligible, at least by typically decreasing surprise. Assuming that science aims to make phenomena as intelligible as possible, we gain a non-practical epistemic reason to prefer formulations that make a given phenomenon more manifest.

A simple example from quantum field theory illustrates these connections between epistemic possibilities, surprise, and intelligibility. The Lagrangian density below initially appears to describe an interacting scalar field ϕ, due to the terms third-order and higher, such as ϕ^3 (Cheung 2017, p. 2):

$$\mathcal{L} = \frac{1}{2}[1 + \lambda_1\phi + \frac{1}{2!}\lambda_2\phi^2 + \frac{1}{3!}\lambda_3\phi^3 + \ldots]\partial_\mu\phi\partial^\mu\phi \tag{6.4.1}$$

Written in this form, the Lagrangian density leaves open the possibility that it describes an interacting field. It is not the case that one ought to know whether the amplitudes that describe scattering n-many particles vanish. We might then go on to calculate the 'tree-level' amplitude for scattering four particles (i.e. to first order in perturbation theory). We would find that it vanishes, reflected by the cancellation of a few Feynman diagrams. Intrigued, we might press on, calculating 5-point amplitudes and higher. We would find that each vanishes. As Cheung notes, "the 14-particle amplitude also vanishes, albeit through the diabolical cancellation of upwards of 5 trillion Feynman diagrams" (2017, p. 3). Well before this point, we might already suspect that the Lagrangian density (6.4.1) actually describes a free scalar field.

Indeed, by performing a suitable field redefinition, we can transform the density (6.4.1) into one that manifestly describes a free scalar field.[11] A manifestly free theory rules

[11] In general, scattering matrix elements are invariant under field redefinitions $\phi \to f(\phi)$ such that f'(0) =1 (Cheung 2017, p. 3). For at least this reason, we ought not take the Lagrangian density or the Feynman rules for vertices too literally!

out the possibility that there are non-vanishing amplitudes describing particle scattering. Hence, the vanishing of these n-point amplitudes becomes unsurprising and to that extent more intelligible. We expect that a free scalar has trivial interactions with itself. In a claim consilient with many themes of this chapter, Cheung notes that "a poor choice of field basis may obscure or *altogether conceal* certain underlying structures of the theory" (2017, p. 3).

A similar moral arises in the context of conjunctive normal form and tautological sentences. Given a structurally complicated or 'concealed' tautology, we might check whether it is a tautology by computing each row of its truth table. As we proceed, we might begin to suspect that we are dealing with a tautology. The truth-value of certain rows might initially seem surprising. By contrast, if we were to convert this sentence into a logically equivalent conjunctive normal form, then its tautological status would be manifest. It would then be unsurprising that each row of its truth table evaluates to true. The possibility of any row evaluating to false would have already been ruled out.

6.4.3 Degrees of manifestness

The connection between i) making a property manifest and ii) ruling out epistemic possibilities suggests a promising strategy for gradating the notion of manifestness. A formulation makes a property manifest to the extent that it rules out possibilities where the property does not obtain. For instance, consider a musician who has 'good but not perfect pitch,' someone who can typically identify a tone to 'plus or minus' the actual pitch-class. Intuitively, the pitch is more manifest to them than to someone who completely lacks a musical ear. On the criterion I am proposing, this is because a musician with good-but-not-perfect pitch rules out more epistemically possible solutions than an ordinary person.

To make this criterion precise, we require a measure on the space of epistemically possible solutions. To determine which of two formulations or variable choices makes a given property more manifest, we must compare the possible solutions that they rule out. Perhaps there is a uniform way of quantifying such epistemic possibilities.[12] Regardless, it seems that we can at least suggest plausible measures in many problem-solving contexts. For instance, when it comes to determining whether a truth-functional sentence

[12]Perhaps such criteria could be constructed from information theory, topology, or truth-maker semantics.

is a tautology, each row of the truth table contributes two epistemic possibilities: true or false under that truth value assignment. A choice of expressive means that rules out more of these possibilities counts as making a given property more manifest.

Moreover, there seem to be good independent reasons for taking seriously the idea of "a space of epistemically possible solutions" for a problem. Generically, we can understand problem-solving as a process of structuring a space of possible solutions. Epistemically-different problem-solving plans result in different structurings of this space: they rule out or in different possibilities. Other things equal, we have epistemic reasons to prefer those problem-solving plans that restrict the space of solutions as much as possible. Applying the account of better understanding from Section 4.5, this means that manifest formulations provide better understanding of the phenomena (i.e. we have a non-practical epistemic reason for preferring a manifest formulation). The same kind of reasoning applies to the use of symmetry groups of differential equations: identifying such groups epistemically constrains the solutions of differential equations that obey those symmetries. This is one of the insights that led Wigner to apply symmetries to quantum mechanics in the 1920s.

At least in some scientific contexts, the space of possible solutions seems highly concrete and far from metaphorical. Physicists provide precise characterizations of such epistemic possibilities whenever they construct a space of possible values for an unknown parameter. Many experimental searches in cosmology and particle physics aim to restrict this space of epistemically possible values as much as possible. Although we may not be able to achieve this level of precision in an arbitrary problem-solving context, it at least supplies a helpful model for philosophical theorizing about reformulations.

6.4.4 Problem-solving adequacy and fruitfulness

In many of the examples from Sections 6.2 and 6.3, the more manifest formulation makes available a problem-solving plan that a less manifest formulation does not support. This is particularly striking in the case of 'sleeve properties.' By wearing a property on the sleeves, the manifest formulation allows us to solve the problem by 'reading off' this property from the expression. For instance, a manifestly Lorentz covariant expression supports a problem-solving plan for Lorentz covariance that a non-manifestly covariant expression does not support (at least without further transformations). Similarly for the case of German vs. English: German makes available a problem-solving plan for guessing the meaning of 'die Speisekarte' that English does not support for 'menu.'

By making alternative problem-solving plans available, these kinds of reformulations contribute to the aim of problem-solving adequacy (introduced in Chapter 5). They supply plans that can succeed in a wider variety of epistemic circumstances. For instance, when it comes to solving the 'menu' problem, a German speaker does not need a dictionary. Likewise, when it comes to determining the pitch-class of a tone, someone with perfect pitch does not require a measuring device.

I conjecture that a more manifest formulation supports alternative problem-solving plans in virtue of ruling out more epistemically possible solutions. Because the English language does not place constraints on the meaning of 'menu' from English subwords, a pure-English speaker has no other recourse than to consult a dictionary (or some other testimonial source). In contrast, German lets us decrease credence in many epistemic possibilities for the meaning of 'die Speisekarte,' such that no dictionary is necessary (at least not necessary for increasing our credence in the meaning of this word).

These differences in problem-solving adequacy amount to differences in fruitfulness. The more manifest formulation supports a plan that can succeed in a wider range of problem-solving contexts, such as contexts where we lack a measuring device for tone or lack a dictionary. Recall that in Chapter 2, I argued that bald appeals to fruitfulness do not provide a satisfying account of the intellectual differences between reformulations. Instead, I urged seeking a local understanding of these differences. Wherever possible, we ought to be able to appraise compatible formulations within a shared domain of problem-solving. Here, we see a local strategy for accounting for differences in fruitfulness: at least some such differences seem to arise from differences in the ruling out of epistemically possible solutions.

6.5 Coordinate Transformations

Coordinate transformations provide one of the simplest cases of philosophically interesting variable changes. Different kinds of coordinate systems sometimes make different properties manifest. Below, I will demonstrate why this matters, using two-dimensional Cartesian vs. polar coordinates as a detailed example. We will see that Cartesian coordinates make manifest the properties of being horizontal or vertical, whereas polar coordinates make manifest the properties of having constant polar angle or constant radius (i.e. being a circle). Other examples of intellectually significant coordinate choices include rectangular vs. spherical vs. cylindrical coordinates in three dimensions, Cartesian vs. internal coordinates in the modeling of molecules, and Eulerian vs. Lagrangian coordinates in fluid dynamics. Although we can express many of the same epistemic dependence relations in these coordinate systems, different coordinate choices nevertheless lead to differences in what we need to know to solve problems.

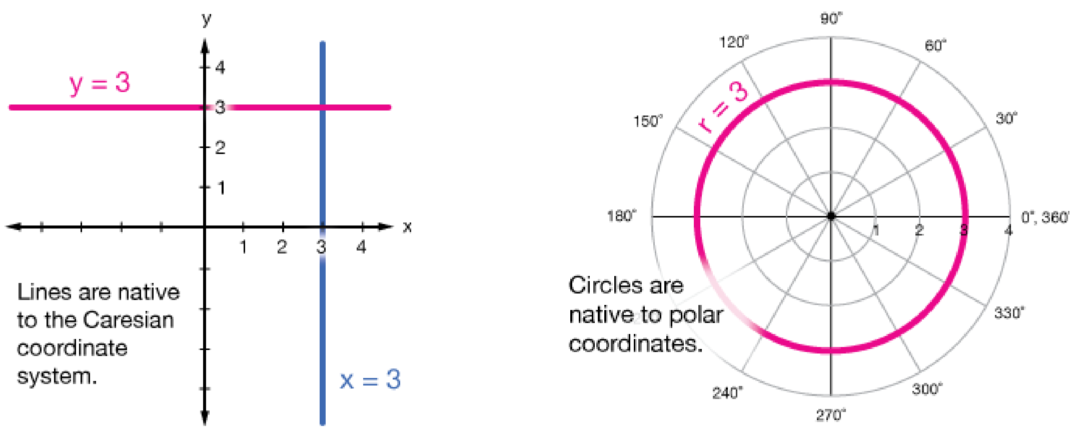
Of course, not all coordinate transformations are intellectually significant. Some coordinate transformations are instead trivial notational variants: they may provide differences in convenience (up to our idiosyncratic conventional preferences), but they evince no intellectually significant differences. As described in Section 2.6, transforming between two Cartesian coordinate systems typically does not provide any differences in EDRs. Such transformations are analogous to systematically replacing every instance of the numeral "5" with "V" in our numeral system. This kind of notational change does not alter what we need to know to solve problems.

One case where coordinate transformations *do* seem to make an intellectual difference is when a system has a symmetry or invariant. For instance, if we are modeling a cylinder, then it is intellectually significant to passively transform to cylindrical coordinates where the z-direction lies along the length direction of the cylinder (so that circular cross-sections of the cylinder are perpendicular to this axis). This makes manifest the length of the cylinder. More precisely, the z-axis now wears the cylinder's length on its sleeves. Such choices amount to a separation of degrees of freedom. Indeed, the examples below involving Cartesian vs. polar coordinates illustrate the same moral. Cartesian and polar coordinates are adapted to equations with different kinds of symmetries or invariant degrees of freedom.

Cartesian vs. Polar Coordinates

Although there is a simple translation procedure between Cartesian and polar coordinates, these coordinate systems are not trivial notational variants. For certain problems, these notations support epistemically different problem-solving procedures. In virtue of these differences in problem-solving plan, Cartesian and polar coordinates make different properties manifest. According to my account of manifest properties, this means that they change when we ought to infer that a system has a given property.

Figure 7 illustrates the different properties that Cartesian and polar coordinates make manifest.[13] These properties define different kinds of graphs. Given a horizontal or vertical line, Cartesian coordinates make manifest the relevant invariant degrees of freedom (the y-coordinate and x-coordinate, respectively). Likewise, given a circle or a diagonal line, polar coordinates make manifest the relevant invariant degrees of freedom (the radius and polar angle, respectively).



(a) Cartesian coordinates make manifest horizontal and vertical lines.

(b) Polar coordinates make manifest circles and diagonal lines, e.g. those through the origin.

Figure 7: Cartesian vs. Polar Coordinates

By *Cartesian coordinates*, I mean the following expressive means: a choice of x and y axes has been made on the plane, with a right angle between them (so that the axes are orthogonal). To represent the equation of a line, we must represent it using these variables x and y. We can imagine working on standard grid paper, representing many distance measurements that we can make using a ruler. By *polar coordinates*, I mean the following expressive means: a choice of reference axis has been made from which to measure the polar angle θ. A choice of origin has been made from which to measure the radial distance r. Again, we can imagine working on polar grid paper, with circles of increasing radii surrounding the origin, and various polar angles indicated with diagonal lines passing through the origin. In both cases, to describe multiple functions at once in a commensurable way, we must keep the reference choices fixed. Hence, in solving the

[13]This image comes from `https://xaktly.com/MathPolarCoordinates.html`.

problems below, it is impermissible to alter the reference choices (e.g. placement of the x-axis, y-axis, angular reference axis, or the origin). This prevents us from trivializing a given problem simply by making a convenient choice of reference axis.

To identify epistemic differences between Cartesian and polar coordinates, I will compare two agents: Carla and Paula. Carla works within a Cartesian coordinate system, whereas Paula works within a polar coordinate system. They are engaged in solving various problems in Euclidean geometry. The philosophical challenge is to locate differences in what these agents need to know at various stages of problem-solving—differences that go beyond the stipulated fact that Carla understands Cartesian coordinates, while Paula understands polar coordinates. If there were no such differences, then Cartesian and polar coordinates would be trivial notational variants after all, in the same way that the English "here is a dog" is synonymous with the German "hier ist ein Hund". Of course, there is trivially an epistemic difference between knowing English and knowing German, but as Section 2.6 describes, that kind of language-dependent epistemic difference does not qualify as intellectually significant.

Here is the first problem: you are presented with a horizontal line drawn in your coordinate system. What is the equation of this line? There are a variety of different ways to proceed, based on different epistemic dependence relations. Using point–slope form, it suffices to know two points on the line, subsequently using these to calculate the slope and intercept of an axis. Alternatively, since the line is horizontal, it suffices to express its vertical displacement from a reference line. Imagine that both Carla and Paula plan to rely on this latter EDR. The question then is whether in executing their plans, any differences arise in which facts are manifest. Specifically, is there a point at which Carla, but not Paula, ought to infer the equation of the horizontal line?

Suppose that Carla and Paula begin in the same way, measuring the vertical displacement of the horizontal line using a ruler (or, perhaps Carla uses the markings on her y-axis, Paula the markings on her r-axis). It turns out that for the given line, the vertical displacement is 5 units from their respective reference lines. At this point, I contend, their problem-solving plans diverge. Since the vertical displacement just is her y-coordinate, Carla ought to infer that the equation of the line is $y = 5$, thereby arriving at the Cartesian solution. In contrast, Paula cannot yet express the equation of the line in her coordinate system, despite knowing that the vertical displacement is 5 units. Paula needs to know something further, namely she needs to know how to express vertical displacement in polar coordinates. Specifically, Paula needs to know that $y = rsin\theta$, relating the height of a right triangle to its hypotenuse and the angle opposite the height. In contrast, Carla did not need to invoke any translation procedure. There is thus a difference in what Carla and Paula need to know, even once we control for language-dependent epistemic differences. Compare the structurally parallel example from Section 6.3: Ender would need to translate 'menu' into German in order to carry out Gertrude's problem-solving plan (which relies on the linguistic substructure of 'die Speisekarte').

This example shows that Cartesian coordinates make manifest the property of being horizontal. A line is horizontal whenever its vertical displacement is invariant. Cartesian coordinates focus attention on the vertical displacement as one of the basic degrees of freedom, namely the coordinate y. They trivially wear vertical displacement on the sleeves. Hence, upon measuring the vertical displacement of a horizontal line, Carla ought to infer the equation of this line in Cartesian coordinates. Since polar coordinates do not focus on the vertical displacement as one of the basic degrees of freedom, it is not the case that Paula ought to infer the equation of the line in polar coordinates. Indeed, it is tempting to make the stronger claim that it would be impermissible for Paula to infer the equation of the line in polar coordinates until she performs this translation. Arguably, Paula needs to know how to express the vertical displacement in polar coordinates. This effectively involves translating from the Cartesian coordinate y to polar coordinates. *Mutatis mutandis*, we see that Cartesian coordinates also make manifest the property of being vertical, i.e. of having invariant horizontal displacement from a reference line.

Polar coordinates make different properties manifest.[14] These include the properties of i) having constant polar angle and ii) having constant radius (being a circle). Imagine that Carla and Paula are presented with a diagonal line passing through the origins of their respective coordinate systems. Both plan to exploit the following epistemic dependence relation: to determine the equation of a diagonal line through the origin, it suffices to measure the angle between it and a given reference line (the x-axis in the case of Carla; the $\theta = 0$ axis in the case of Paula). To keep things as epistemically symmetric as possible, suppose that both use a protractor to measure the angle, determining that it is 45 degrees.

[14]Such properties are not always mutually exclusive. Both Cartesian and polar coordinates make manifest the defining property of a vertical line through the origin: it has both zero horizontal displacement and constant polar angle 90 degrees.

At this point, Paula ought to infer that the equation of the line is $\theta = 45^\circ$. The equation of the line is already manifest to her.

In contrast, Carla is not yet permitted to infer the equation of the line in Cartesian coordinates (namely, the fact that $y = x$). Instead, she needs to know a further fact, namely how to relate this reference angle of 45 degrees to an expression involving the Cartesian coordinates x and y. Carla has effectively measured the polar angle θ, and she needs to know how to translate this angular degree of freedom into Cartesian coordinates. Specifically, she needs to know that $\theta = arctan(y/x)$. From this equation, she can infer that $y/x = tan(\theta) = tan(45^\circ) = 1$. After this series of inferences, Carla ought to infer that $y = x$, thereby solving the problem in Cartesian coordinates. Although Carla exploited the same initial plan as Paula—namely, the directive to measure the angle that the line makes with a reference line passing through the origin—she required additional knowledge to solve the problem, knowledge that Paula did not require in polar coordinates.

Mutatis mutandis, the same lesson applies to circles centered at the origin. Since these geometric objects have constant radii, polar coordinates make their equations manifest. For instance, upon measuring the radius of such a circle to be 5 units, Paula ought to immediately infer that its equation is $r = 5$. In contrast, Carla needs to know how to relate this radius to Cartesian coordinates, using the trigonometric fact that $r = \sqrt{x^2 + y^2}$.

These examples evince subtle epistemic differences in the choice of expressive means. To appreciate them, it may help to recall the case of a person with perfect pitch. When it comes to horizontal and vertical lines, Carla is like someone with 'perfect pitch' for these. Upon a minimal measurement (analogous to hearing the pitch), she ought to immediately infer the equation of the line. Likewise, Paula has 'perfect pitch' for diagonal lines through the origin and circles centered at the origin. Upon measuring the polar angle or radius, Paula ought to immediately infer the equations for these kinds of geometric objects. In contrast, Carla is like someone who lacks perfect pitch for these geometric objects: she has to do further inferential work in order to determine their equations.

As a final and perhaps more dramatic example, consider Archimedean spirals. Polar coordinates make manifest the defining property of Archimedean spirals: the radius increases as a constant proportion of the polar angle, i.e. $r = a + b\theta$, for some constants a and b. Cartesian coordinates obscure this property. In the simplest case where $r = \theta$, the corresponding Cartesian equation is $y = x\,tan(\sqrt{x^2 + y^2})$. This equation points to another interesting epistemic difference: in polar coordinates, it is possible to characterize r explicitly in terms of the polar angle θ. Yet, it is (seemingly) not possible to characterize y explicitly in terms of the Cartesian coordinate x. Instead, the best we can do is represent the graph of the Archimedean spiral implicitly in Cartesian coordinates.[15]

6.6 Preferences, Fundamentality, and Privileging

Thus far, I have analyzed the notion of "manifest properties" in terms of what we ought to infer in a particular epistemic circumstance. Reformulating can change our epistemic circumstance, thereby changing what properties are manifest. Nevertheless, some might worry that my account does not go far enough to capture the significance of reformulations that make properties manifest. It is common for scientists and mathematicians to think that one formulation is *more fundamental* than another, but it is unclear how fundamentality could reduce to the epistemic differences that conceptualism focuses on. On this basis, a fundamentalist might claim that conceptualism owes us an account of common judgments of fundamentality. In keeping with the empiricist scruples of Chapter 2, conceptualism must provide an account of fundamentality that avoids metaphysically substantial commitments. To meet this demand, I will provide a non-metaphysical account using resources from metaethical expressivism.

Recapping Expressivism

Metaphysicians and philosophers of science typically assume that declarative sentences about the world should be interpreted as playing a representational role. Expressivism rejects this assumption, observing that "not everything we think or say need be understood as representing the world as being some way" (Brandom 2011, p. 11). As Carnap wrote in 1934, "We have here to distinguish two functions of language, which we may call the expressive function and the representative function" (1935, p. 27).[16] Hoping to elimi-

[15]Note that we could perform another variable transformation to an (r, θ) phase space where r and θ are orthogonal. In this space, the simplest Archimedian spiral $r = \theta$ is characterized by an invariant ϕ, which corresponds to the angle measured from the θ axis. This further choice of variables makes this invariant property of Archimedean spirals *even more* manifest. In this parameterization, we effectively "unroll" the Archimedean spiral into a straight line passing at 45 degrees through the origin; we linearize the graph.

[16]See also Sellars (1958, p. 282), who denies that "the business of all non-logical concepts is to describe." Unlike Carnap, Sellars draws a more egalitarian moral, noting that "many expressions which empiricists have relegated to second-class citizenship in discourse are not *inferior*, just *different*." Carnap's 1934 lectures

nate metaphysics from analytic philosophy, Carnap proceeded to claim that "metaphysical propositions—like lyrical verses—have only an expressive function, but no representative function….They express not so much temporary feelings as permanent emotional or volitional dispositions" (1935, p. 29). With Carnap, I agree that metaphysical statements play an expressive role. However, unlike Carnap, I am agnostic on whether this is the only role that metaphysical statements play. I will argue that we can at least make sense of physicists' judgments of fundamentality as playing a particular expressive role, regardless of whether they play a representational role as well.

Expressivism is a kind of philosophical naturalism: it explicates otherwise puzzling vocabularies in terms of non-mysterious, naturalistically acceptable ones (Price 2011). In this case, I will argue that we can understand physicists' talk about fundamentality in terms of their attitudes toward privileging some formulations or variable choices over others. Some philosophers may nevertheless hanker after something more than this kind of anthropological analysis. Namely, they may desire a representational or descriptive analysis of judgments of fundamentality. I am not inclined to stop them, although I will resist if they contend that I ought to hanker after something more as well. Brandom phrases this resistance to representationalism rather eloquently:

> If the practices themselves are all in order from a naturalistic point of view, any difficulties we might have in specifying the kind of things those engaged in the practices are talking about, how they are representing the world as being, ought to be laid at the feet of a Procrustean semantic paradigm that insists that the only model for understanding meaningfulness is a representational one. (Brandom 2011, p. 192).

In Section 4.5, I provided an expressivist account of comparative judgments of understanding. I argued that we can understand judgments of the form "X provides better understanding than Y" as expressing a mental state of *being for intellectually-preferring X to Y*. Equivalently, when someone judges a formulation X to provide better understanding than Y (for a particular problem), they endorse a set of norms that permit intellectually-preferring formulation X to formulation Y (at least for this kind of problem). In this way, we can vindicate scientists' and mathematicians' ordinary judgments about comparative understanding without having to posit metaphysically substantial facts or properties about comparative intellectual value. Structurally, this parallels how metaethical

have been reprinted in Carnap (1996 [1935]).

expressivists aim to vindicate ordinary moral (or normative) judgments without positing metaphysically substantial facts or properties about moral rightness or wrongness (or primitive ought-claims, in the case of normativity) (Blackburn 1998; Gibbard 2003).

Expressivism about Fundamentality

Here, I propose a similar expressivist analysis of fundamentality. Indeed, there is a close connection between judgments of fundamentality and comparative judgments of understanding. To judge that X is more fundamental than Y typically entails that X provides a better understanding of some class of problems or phenomena than Y. As noted in Section 4.5, this judgment of better understanding might be aim-relative. For instance, understanding the human heart as a collection of molecules provides a better understanding relative to certain aims, but not all. A molecular understanding of the heart obscures the mechanical understanding we might achieve by describing the heart at a higher length scale, focusing on biological tissue. Still, there is a sense in which the molecular understanding is more fundamental than the biological understanding.

On my proposal, comparative judgments of fundamentality express an attitude of *being for privileging*. To judge that X is more fundamental than Y expresses a mental state of *being for privileging X to Y*. Privileging is a particularly committal form of preference. In contrast to preferring something, to privilege something typically entails that it is to be *uniquely* preferred, at least along a certain dimension. Privilege is a kind of *maximal preference*: to privilege something is to believe it is uniquely best in some regard. Whereas judgments of comparative or relative fundamentality involve privileging X over Y, we can provide a similar analysis of *absolute fundamentality*. To judge that X is absolutely fundamental is to express a mental state of being for privileging X, relative to all alternatives.

Often, when physicists and mathematicians call a fact—or entity, structure, principle, etc.—'fundamental,' they express an attitude of privileging that fact in *derivations* of other facts. Other things equal, a fact X is more fundamental than a fact Y when X figures in a derivation of Y (but not vice versa). This dimension of fundamentality is intricately connected to metaphysical notions of grounding and truth-making. At first glance, such connections might seem to pose a problem for expressivism about fundamentality. Fortunately, Barker has developed a promising expressivist approach to truth-making claims. To say that some fact(s) X (non-causally) makes it the case that Y is to express a commitment to using X to derive Y (2012, p. 273). Hence, I am optimistic that we can understand this pervasive dimension of fundamentality in terms of a pro-attitude of privileging some facts for certain derivational roles.

Overall, my expressivist analysis of fundamentality provides a simple response to the objection raised above. Conceptualism can sanction scientists' and mathematicians' ordinary judgments of relative and absolute fundamentality. Consider a physicist who claims that variables that makes the hidden hyperspherical symmetry of hydrogen manifest are more fundamental than variables that obscure this symmetry. This judgment of relative fundamentality amounts to endorsing a set of norms that permit privileging the manifest variable choice to the non-manifest variables. As we will see in Sections 6.8 and 6.9, there are a variety of reasons to privilege variables that make a symmetry manifest. Hence, we can endorse these judgments of fundamentality as a rational aspect of scientific and mathematical practice. These judgments play an important functional role in coordinating scientific and mathematical problem-solving. They help scientists converge on variable choices that have instrumental and epistemic value. We can endorse these judgments without committing ourselves to metaphysically substantial facts or properties about fundamentality. Instead, we simply focus on the non-descriptive functional roles that judgments of fundamentality perform.

Fundamentality and Invariants

In problem solving, we are often interested in invariant properties. Such properties allow us to characterize systems or objects across varying contexts. They provide a stable point of reference. Hence, scientists and mathematicians have a good epistemic reason to prefer expressive means that make an invariant property manifest. Expressive means that wear an invariant on their sleeves are better suited to make invariance manifest. In some sense, they *minimize* what we need to know to determine invariance. This is perhaps one reason why we often associate invariant degrees of freedom with more fundamental properties. Additionally, in physics, observables must be invariant under the symmetries of a theory. Expressive means that obscure these invariances are therefore rightly viewed as less fundamental, *ceteris paribus*: we have at least an epistemic reason to disprefer them.

In general, scientific preferences for variable choices might align with the following methodological advice: if one plans to use an epistemic dependence relation that involves a particular degree of freedom, then it is better to express that EDR in a notation that trivially wears that degree of freedom on the sleeves (i.e. is manifest to the 0th degree). Doing so will typically make a property of interest manifest, namely the property that we are using the EDR to assess. Accepting this methodological advice does not involve any further metaphysical commitments to whether this degree of freedom is fundamental in any deep sense.

Moreover, to say that a notation is particularly well-suited for expressing a particular EDR is not to say that it is uniquely suited. There could be a wide variety of expressive means that are equally well-suited for making a particular property manifest. Hence, on the account I defend, we do not have to view scientists as aiming for a single, overarching, most fundamental language. Instead, we can interpret their judgments of fundamentality as often being implicitly relativized: X is more fundamental than Y relative to a certain class of problems or a certain set of aims.

My expressivist account of fundamentality does not preclude a descriptivist or representationalist account. For all I say here, some such account could be correct. I simply claim the following: regardless of whether physicists's judgments of fundamentality amount to anything more, they at least play the functional roles that my expressivist account describes. For my purposes, it is enough to vindicate physicists's ordinary judgments of fundamentality. Unlike Carnap, I do not intend to rule out or eliminate substantial metaphysics. I have a weaker aim, namely to show that many of us can responsibly go on without such metaphysics. As Brandom notes, "a successful local expressivism about some vocabulary [e.g. fundamentality] would show that, while it might be *possible* to offer a representational semantics for that vocabulary, it is not *necessary* to do so in order to show it to be [naturalistically] legitimate" (2011, p. 195).

An Objection from Instrumentalism

An instrumentalist about reformulations (see Section 2.4) might object to my account of fundamentality as follows: sometimes, our overall reasons for privileging a choice of variables contains a confluence of epistemic and practical values. For instance, even though polar coordinates make manifest the invariant polar angle of a diagonal line, we might still prefer to use Cartesian coordinates to determine the equation for this line. We might prefer Cartesian coordinates for a variety of practical or idiosyncratic reasons: perhaps we dislike polar coordinates in general, or do not have a protractor, or prefer to always find the equations of straight lines using point–slope form (since this works for any straight line), etc. Surely—the instrumentalist objection continues—these reasons for privileging Cartesian coordinates have nothing to do with fundamentality.

To respond to this instrumentalist objection, it suffices to note the following: expressivism does not rely on a dispositional account of our attitudes or preferences. Instead, it relies on a fitting-attitudes account. Certain reasons are fitting for particular attitudes. For instance, an expressivist about humor does not say that jokes are funny because people laugh at them. Instead, jokes are funny when people *ought* to laugh at them. There are a wide variety of non-humor related reasons why someone might laugh at a joke. Expressivism can rightly classify those non-humor-related reasons as irrelevant to the comedic value of the joke.

Similarly, even if we dislike polar coordinates, we can still recognize that they make manifest the invariant degrees of freedom of a number of different kinds of equations. We can recognize that this gives us a reason for viewing polar coordinates as more fundamental than Cartesian coordinates for describing such equations. In other words, we recognize that reasons of personal preference are not the right kinds of reasons for judgments of fundamentality. They are not fitting to this end. Hence, an expressivist about fundamentality can agree with the instrumentalist that we often prefer certain variables for instrumental or idiosyncratic reasons. All the while, we can recognize that these instrumental reasons are not the right kinds of reasons for viewing one formulation as more fundamental than another.

6.7 Gauge Choices

In Lagrangian quantum field theory, gauge choices provide an illuminating example of how different formulations can make different properties manifest. In particular, different gauge choices illustrate trade-offs that can arise between different formulations. As we have already seen in the simpler context of polar vs. Cartesian coordinates, making one property manifest can come at the cost of obscuring others.

Indeed, one such trade-off arises whenever we introduce gauge degrees of freedom in the first place. On physical grounds, we know that a massless gauge field with non-zero spin—such as the photon field—has only two degrees of freedom (two physical polarization states). Nevertheless, in order to write the Lagrangian density (and hence the action) in a manifestly Lorentz invariant form, we introduce two redundant, gauge degrees of freedom. These allow us to write the gauge field A^{μ} as a 4-vector, supporting our syntactic criteria for manifest Lorentz covariance. As Cheung puts it, "these redundant modes are a necessary evil of manifest Lorentz covariance" (2017, p. 2). Hence, we trade-off manifest physical degrees of freedom for manifest Lorentz invariance. Why do physicists so often make this trade? By enforcing Lorentz invariance in the Lagrangian density, we massively constrain the space of possible interaction terms. This strategy has tremendous epistemic power for theory construction.

Here, I will focus on comparing two families of gauge choices: i) manifestly Lorentz covariant gauges vs. ii) manifestly unitary gauges. As their names indicate, they respectively make manifest the properties of Lorentz covariance and unitarity. They also each obscure the property that the other makes manifest, illustrating a trade-off. We can understand these gauges as having a symbiotic relationship: to prove that a quantum field theory is unitary, it is best to use a manifestly unitary gauge. In contrast, for most other calculations, it is best to use a manifestly Lorentz covariant gauge, since they tend to simplify calculations (Siegel 2005, p. 30). Fortunately, since these are compatible formulations, we are not forced to choose between them. Gauge choices like these provide evidence that we can understand particle physicists as exploiting different variable choices in different contexts, rather than as aiming at a single fundamental language for describing scattering processes. As Siegel notes, we have "different gauges for different uses" (2005, p. 13). Against Maudlin's (2018, pp. 14, 16) methodological recommendations, I deny any need to interpret these gauge choices as leading to competing or rival ontologies.[17]

Before delving into these gauge choices, a few remarks on "unitarity," i.e. the property of being unitary. A quantum field theory is unitary provided that it satisfies two conditions: i) all probabilities for scattering processes are non-negative and ii) probability is conserved, i.e. the probabilities of all possible processes sum to one. This second con-

[17]It is only in contexts where we view the gauge field A^{μ} as being a calculational device—such as some interpretations of classical electromagnetism—that Maudlin sanctions interpreting different gauge choices as leading to compatible formulations. By also interpreting different gauge choices in quantum field theory as leading to compatible formulations, I violate Maudlin's interpretive norms. Maudlin may view me as being afflicted with "the attitude of the engineer rather than the natural philosopher" (2018, p. 6). So much the worse for the natural philosopher, say I!

dition amounts to the Hamiltonian being Hermitian, i.e. $H^\dagger = H$ (Siegel 2005, p. 355).[18] The first condition requires that the inner product on Hilbert space is positive definite. This condition is more difficult to check, and it is the one that unitary gauges help make manifest.

Manifestly Lorentz Covariant Gauges

Lorenz gauge provides a constraint on the gauge field A^μ that is manifestly Lorentz covariant: $\partial_\mu A^\mu = 0$. By constraining the gauge field in a manifestly covariant manner, we preserve the manifest covariance of those expressions that were already manifestly covariant before we imposed this constraint.

In Lagrangian quantum field theory, we generalize Lorenz gauge to the family of R_ξ gauges. To gauge-fix in this manner, we add a manifestly Lorentz invariant term to the Lagrangian: $-\frac{(\partial_\mu A^\mu)^2}{2\xi}$. Different values of the parameter ξ result in different gauge-fixings. Provided that the Lagrangian is already manifestly Lorentz invariant, the additional R_ξ term preserves this manifest Lorentz invariance.

Two common R_ξ gauges are Landau gauge and Feynman–'t Hooft gauge, which set ξ equal to zero and one, respectively. Landau gauge recovers Lorenz gauge in the limit as ξ goes to zero. Feynman–'t Hooft gauge ($\xi = 1$) is particularly advantageous for explicit calculations because it tends to give the simplest form for the propagator terms. In general, propagators in R_ξ gauge take the form $2[\frac{\eta_{ab}}{p^2} + (\xi - 1)\frac{p_a p_b}{(p^4)}]$ (Siegel 2005, p. 389). Clearly, setting $\xi = 1$ results in the simplest propagator term: $2\frac{\eta_{ab}}{p^2}$. These gauges also have the advantage of easily generalizing from Abelian to non-Abelian symmetry groups.

According to Siegel, the R_ξ gauges "manifest as many global invariances as possible" (2005, p. 386). By preserving manifest Lorentz covariance, the R_ξ gauges trivialize the preservation of these space-time symmetries. In other words, it becomes unnecessary to explicitly calculate that these symmetries are preserved. Instead, the expressions continue to wear these properties on the sleeves. Wearing properties on the sleeve has non-practical epistemic value (in addition to any practical value it might have as well). The symmetry properties of these expressions become more intelligible, at least on account of becoming less surprising.

[18] Alternatively, a unitary quantum field theory has a unitary evolution operator U, where this means that $U(t_2,t_1)^\dagger U(t_2,t_1) = I$. This requirement amounts to the conservation of probabilities (Siegel 2005, p. 298).

Manifestly Unitary Gauges

I turn now to manifestly unitary gauges. These include light cone and space cone gauge. Not only do these gauges make unitarity manifest, but also they eliminate unphysical degrees of freedom (such as those coming from ghost fields). I will focus in particular on light cone gauge, which according to Siegel is "the simplest for analyzing physical degrees of freedom, since the maximum number of degrees of freedom is eliminated" (2005, p. 211).

Light cone gauge relies on a light cone basis, which uses a different basis for the metric η^{ab}. Rather than focus on the A^0 and A^1 components of the gauge field A^μ, we focus on their linear combinations, calling the resulting components A^+ and A^-, where $A^\pm = \frac{1}{\sqrt{2}}(A^0 \pm A^1)$. To work in the light cone gauge, we first fix one degree of freedom by setting $A^+ = 0$. To eliminate the second gauge degree of freedom, we introduce the component A^- as an auxiliary field in the Lagrangian density $\mathcal{L}$, ultimately eliminating it (Siegel 2005, p. 210). We thereby reduce the four degrees of freedom in A^μ to two, representing the actual physical degrees of freedom of the gauge field.

As mentioned above, unitarity requires that the inner product on Hilbert space be positive definite (this amounts to a requirement that the energy is positive). The sign of the energy is intimately connected with the sign of the kinetic term in the Lagrangian density. By eliminating unphysical degrees of freedom, light cone gauge sets up a simple correspondence between the sign of the kinetic terms and unitarity. Hence, one can 'read off' unitarity from the Lagrangian density when it is written in light cone gauge. We simply require that boson fields have a negative kinetic energy term while fermion fields have a positive one (Siegel 2005, p. 357). In this way, the Lagrangian density in light cone gauge wears unitarity on the sleeves, thereby making it manifest.

Trade-offs and Fundamentality

These two families of gauge choices illustrate the kinds of trade-offs that frequently arise when we change variables. On the one hand, manifestly Lorentz covariant gauges make manifest a (contextually) fundamental symmetry. Nevertheless, they obscure both unitarity and some physical degrees of freedom. On the other hand, manifestly unitary gauges obscure Lorentz invariance, despite eliminating a greater number of unphysical degrees of freedom. A fundamentalist might be inclined to weigh these trade-offs in an attempt to determine which choice of variables is more fundamental *tout court*, or which leads to a more virtuous physical theory (perhaps Maudlin (2018, p. 20) would make this recommendation). I am pessimistic about the prospects of this approach. Clearly, we can use both gauge choices in different contexts. Insofar as a physicist might be inclined to say that one gauge choice is more fundamental in a particular context, we can understand them as expressing an attitude of being for privileging this gauge choice in such contexts (Section 6.6).

Perhaps a fundamentalist might reason as follows: it is epistemically possible for there to be a variable choice that makes manifest all of these properties, with none of the drawbacks. Such a choice of variables or gauge would make manifest i) Lorentz covariance, ii) unitarity, and iii) eliminate unphysical degrees of freedom. If we had such a choice, it would be more fundamental than either of the gauge choices discussed above. Perhaps then, physicists or metaphysicians should be aiming for such a choice of variables. In many ways, spinor–helicity variables accomplish some of these aims. Yet, they also introduce trade-offs of their own. In particular, spinor–helicity variables i) obscure the property of locality and ii) introduce unphysical complex momenta (Elvang and Huang 2015, p. 61). This provides grounds for pessimism that physics will in general arrive at a choice of variables that make manifest all fundamental properties. At least sometimes, when we make one physically significant property manifest, it comes at the cost of obscuring others.[19] Of course, I have looked at only a small set of cases. Nevertheless, these examples motivate a more extensive inductive argument (for future work) that would parallel the Pessimistic Meta-Induction against scientific realism.

6.8 Manifest vs. Hidden Symmetries of Hydrogen

The symmetries of the hydrogen atom provide a striking contrast between manifest vs. hidden properties. In elementary presentations, the hydrogen atom has a manifest spherical symmetry but a hidden hyperspherical symmetry. My account of manifest vs. hidden properties from Section 6.2 makes these claims precise. There is an epistemic circumstance where i) one ought to infer that hydrogen has spherical symmetry but where ii) it seems impermissible to infer that hydrogen has a larger hyperspherical symmetry (at least in this

[19] I thank Henriette Elvang for encouraging me to weaken some more sweeping claims in favor of pessimism.

epistemic circumstance). Before making this further inference, more inferential work is required: one must transform the given epistemic circumstance into a different one.

Many models exist for the hydrogen atom, but not all of them exhibit a hidden hyperspherical symmetry. Hence, I will focus on a model for a nonrelativistic, spinless hydrogen atom. This model was the first to be worked out after the advent of quantum mechanics (Pauli 1926). Despite neglecting relativity and electron spin, this simple model provides a robust first-order approximation of the hydrogen atom's energy-level spectrum.[20]

Manifest Spherical Symmetry

In nonrelativistic quantum mechanics, we determine properties of a system by analyzing its Hamiltonian, often in conjunction with the Schrödinger equation. For a nonrelativistic, spinless hydrogen atom, the Hamiltonian consists of two terms: a kinetic term for a free particle and a potential energy term given by Coulomb's law of electrostatics:[21]

$$H = \frac{\mathbf{p}^2}{2\mu} + V(x,y,z) = -\frac{h^2}{8\pi^2\mu}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r} \tag{6.8.1}$$

On its own, the Hamiltonian (6.8.1) makes manifest that the hydrogen atom has spherical symmetry. This is because both the kinetic and potential terms are manifestly invariant under arbitrary rotations in three-dimensional Euclidean space, entailing that H is likewise spherically symmetric (since a sum of spherically symmetric terms is spherically symmetric). Clearly, the various constant terms in the expression are invariant under three-dimensional rotations, so all we need to do is check the invariance of the non-constant functions, namely ∇^2 and $1/r$. To see that the kinetic term is spherically symmetric, it suffices to unpack the ∇^2 operator, known as the Laplacian: $\nabla^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2}$. With each Cartesian coordinate on equal footing, this term is invariant under three-dimensional rotations. Turning to the potential term, the function $1/r = 1/\sqrt{x^2+y^2+z^2}$ again places each of the three Cartesian coordinates on equal footing, so its rotational invariance is manifest. Since each term is rotationally invariant, so is the Hamiltonian. It thus has at least the symmetry of the group of proper rotations in three-dimensional

[20] More sophisticated treatments using the Dirac equation and quantum electrodynamics later accounted for higher-order features of the hydrogen spectrum. However, they break the special "dynamical" symmetry of this simple model.

[21] Here, μ is the reduced electron mass $\frac{m_e m_p}{m_e + m_p}$, a function of the electron and proton masses.

Euclidean space, known as the *special orthogonal group in three-dimensions*, $SO(3)$.

Based on the account in Section 6.2, to say that the Hamiltonian H has manifest spherical symmetry is to say that we ought to infer that it is spherically symmetric. More precisely, provided that we implement the problem-solving plan described above, we are in an epistemic circumstance where we ought to infer $SO(3)$ symmetry. Since hydrogen has a spherically symmetric Hamiltonian, it follows that the hydrogen atom has at least this symmetry.[22] If we do not make this inference based on the reasoning above, then we have made an epistemic mistake. We would be doing something epistemically deficient.

Indeed, the foregoing analysis shows that we can say something even stronger: the Hamiltonian in (6.8.1) *wears its spherical symmetry on the sleeves*. The property is not only manifest, but it is made manifest solely on the basis of features of equation (6.8.1) that are already manifest, i.e. manifest before we implement the problem-solving plan detailed above. These already-manifest properties include the placement and identity of the various terms in the expression. On the basis of these syntactical properties, we ought to infer that the constant terms, ∇^2, and $1/r$ are all spherically invariant.[23] On the basis of these inferences, we then ought to infer that H is spherically invariant as well. $SO(3)$ symmetry is a sleeve property of the hydrogen atom Hamiltonian.

Hidden Hyperspherical Symmetry

We can now contrast the manifest status of $SO(3)$ symmetry with the completely different epistemic situation for hydrogen's hidden symmetry. This hidden symmetry is associated with special features of the two-body problem with a $1/r$-potential, leading many physicists to deem it a "dynamical symmetry"—in contrast with "geometrical symmetries" that arise from spacetime symmetries.[24] It turns out that this simple model of the hydrogen

[22] In this context, the symmetry group of a system is defined as the group of operators that commute with its Hamiltonian. Since the Hamiltonian is invariant under three-dimensional rotations, all of these operators commute with H, i.e. $[H, R] = HR - RH = H - H = 0$, for any $R \in SO(3)$.

[23] As shown above, the spherical invariance of these terms becomes manifest when we implement a problem-solving plan for them, such as writing out '$1/r$' explicitly as a function of Cartesian coordinates. This example thereby illustrates the gradated nature of manifest properties.

[24] In this context, a dynamical symmetry refers to a symmetry that is associated with the particular form and nature of the dynamics, e.g. the particular form of a force law, number of interacting subsystems, or energy state of the system (bound or scattering). Note that this is a narrower notion of "dynamical symmetry" than that commonly found in the philosophy of physics literature, where *dynamical symmetries* are those that leave the model's equations of motion invariant. In this broader sense of dynamical symmetry, hydrogen's $SO(3)$ symmetry is also dynamical.

atom has a much larger symmetry group, namely the symmetry of a four-dimensional Euclidean hypersphere. Formally, this group is known as the special orthogonal group in four dimensions, denoted by '$SO(4)$.'

Following the account of Section 6.2, we can at least say that this hyperspherical symmetry is *not manifest*: when presented merely with the Hamiltonian in equation (6.8.1), it is not the case that we ought to infer that this Hamiltonian has hyperspherical symmetry. We do not make an epistemic mistake if we fail to make this inference. Thus, there is an epistemic circumstance C where i) we ought to infer that hydrogen has $SO(3)$ symmetry, but ii) it is not the case that we ought to infer that it has $SO(4)$ symmetry.

Indeed, I am tempted to assert a stronger claim: not only is the hyperspherical symmetry not manifest in this epistemic circumstance, it is hidden. In other words, if we were to infer that H has hyperspherical symmetry solely on the basis of this epistemic circumstance, then we would make an epistemically impermissible inference. We would be jumping to conclusions in an irrational manner.[25] In order to license the inference that hydrogen has hyperspherical symmetry, more epistemic work is required. We must transform our epistemic circumstance into one where we are rationally permitted to infer this symmetry.

It is precisely this transformation of epistemic circumstances that Fock undertook in his analysis of the hydrogen atom (1935b).[26] By changing variables to momentum space, Fock was able to make manifest the hyperspherical symmetry of hydrogen. Schematically, Fock's argument proceeds as follows: we write the integral form of the Schrödinger equation in momentum space. Using a stereographic projection from the three-sphere S^3 to Euclidean three-space $\mathbb{R}^3$, we then demonstrate that this equation is equivalent to an integral equation for the four-dimensional spherical harmonics. We thereby see that the four-dimensional spherical harmonics are solutions to the hydrogen atom's Schrödinger equation. Since these spherical harmonics have $SO(4)$ symmetry, so must the hydrogen atom. Hence, by the end of this argument, the hyperspherical symmetry of hydrogen has become manifest (although it is plausibly not worn on the sleeves of a corresponding expression). This schematic discussion suffices for my philosophical aims here. For the

[25]Compare Field's (2018, p. 5) discussion (stemming from Boghossian) of someone applying an inference rule that—although sound—has not yet been demonstrated to be sound. Such a person would plausibly strike us as being irrational, even if their inference follows a reliable pattern.

[26]See Fock (2005) for an English translation. See also Fock (1935a).

interested reader, I provide more details about Fock's argument below. Less interested readers can happily skip ahead to Section 6.9.

Fock's Argument, in more Detail

Using the Hamiltonian in equation (6.8.1), we can write the time-independent Schrödinger equation $H\psi = E\psi$ for the hydrogen atom, where E is an energy eigenvalue of the hydrogen wavefunction ψ. This results in the following equation, written in position space:

$$-\frac{h^2}{8\pi^2\mu}\nabla^2\psi(x,y,z) - \frac{e^2/4\pi\varepsilon_0}{\sqrt{x^2+y^2+z^2}}\psi(x,y,z) = E\psi(x,y,z) \tag{6.8.2}$$

Fock performs a Fourier transform on this Schrödinger equation, expressing it in momentum space:

$$\frac{\hbar^2}{2m}|p|^2\psi(p) - e^2\sqrt{\frac{2}{\pi}}\int_{\mathbb{R}^3}\frac{\psi(p')dp'}{|p-p'|^2} = E\psi(p) \tag{6.8.3}$$

The form of this equation motivated Fock to consider a stereographic projection from S^3 to $\mathbb{R}^3$. According to McIntosh, "In this form, the kernel can be recognized as the Jacobian determinant for a stereographic projection from the surface of a four-dimensional sphere to three dimensions, which in turn suggests writing the Schrödinger equation in terms of angular variables on the hyperspherical surface" (1971, p. 81). Fock denotes these angular variables as (α,θ,ϕ), and introduces a function $\Psi(\alpha,\theta,\phi)$ defined on the hypersphere. $\Psi(\alpha,\theta,\phi)$ depends as well on the momentum and energy of the atomic state. Using this function, he expresses the Schrödinger equation on the hyperspherical surface as follows:

$$\Psi(\alpha,\theta,\phi) = \frac{\lambda}{2\pi^2}\int\frac{\Psi(\alpha',\theta',\phi')d\Omega'}{4sin^2(\omega/2)} \tag{6.8.4}$$

Here, $\lambda = \frac{me^2}{h\sqrt{-2mE}}$ and $d\Omega$ is the surface element for the 3-sphere. The term $4sin^2(\omega/2)$ in the denominator of the integrand represents the square of the distance between the two points (α,θ,ϕ) and (α',θ',ϕ') on the 3-sphere (hence, ω is the arclength of the great circle that connects the two points) (Fock 2005, p. 288).

Fock then compares this reformulation of the Schrödinger equation to the integral equation for the four-dimensional spherical harmonics:

$$r^{n-1}\Psi_n(\alpha,\theta,\phi) = \frac{n}{2\pi^2}\int \frac{\Psi_n(\alpha',\theta',\phi')d\Omega'}{1-2rcos(\omega)+r^2} \tag{6.8.5}$$

Setting $\lambda = n$ and $r = 1$, we recover the same form as the Schrödinger equation (6.8.4). McIntosh interprets this case as being "the Poisson kernel for a hyperspherical surface harmonic in the degenerate case in which the field point has fallen onto the surface," where $r = 1$ specifies the surface (1971, p. 81). Physically, the integer n is the principal quantum number, labeling the hydrogen atom's energy levels. Due to this correspondence between the two equations, we see that the hydrogen atom wavefunctions can be expressed in terms of the hyperspherical harmonics. Hence, any symmetry of these harmonics is a symmetry of the hydrogen atom wavefunctions, and thus of the hydrogen atom itself.[27]

Fock summarizes the conclusion of his argument as follows:

> Thus we have shown that the Schrödinger equation (6.8.3) or (6.8.4) can be solved with four-dimensional spherical harmonic functions. At the same time the transformation group of the Schrödinger equation has been found: this group is obviously identical to the four-dimensional rotation group. (Fock 2005, p. 289)

Alternatively, we can interpret Fock as having constructed a representation of the group $SO(4)$ on the phase space of the hydrogen atom (namely, the space of square integrable functions on $\mathbb{R}^3$). Fock implicitly shows that this representation commutes with the Hamiltonian for hydrogen. This entails that the hydrogen atom has hyperspherical symmetry (Singer 2005, p. 283).

More precisely, this symmetry applies only to *bound states* of hydrogen, namely those where the electron has negative potential energy. These states constitute the discrete or 'point' spectum for hydrogen. If the electron acquires enough energy, it enters a scattering state (positive potential energy), leading to a continuous spectrum. In this case, the symmetry is that of the Lorentz group, and the geometrical interpretation relies on a hyperboloid rather than a hypersphere (McIntosh 1971, p. 81; Fock 2005, p. 292).

Note that Fock's momentum space representation (6.8.3) of the hydrogen atom Schrödinger equation plausibly does *not* wear the hyperspherical symmetry on its sleeves. Hence, although we have made the symmetry manifest by the end of the

[27]Note that the four-dimensional hyperspherical harmonics have hyperspherical symmetry in virtue of being the angular part of the solutions to the Laplace equation in four dimensions.

derivation (one ought to infer that the system has the symmetry), we have not done so by making the symmetry into a sleeve property of a corresponding expression.

6.9 Hidden Symmetries in $\mathcal{N} = 4$ super Yang–Mills Theory

Precision calculations for predictions at the Large Hadron Collider increasingly require calculations at third-order or higher in perturbation theory. These calculations are necessary to gain a better theoretical understanding of background processes. Without theoretical knowledge of the background, it is impossible to isolate new physics from already understood processes. This task is challenging largely because of how quickly the number of terms grows in perturbation theory. To manage this computational complexity, physicists have had to repeatedly reformulate their calculational techniques. Feynman diagrams provide one such reformulation, but these techniques become infeasible for scattering more than a few particles, due to the rapid growth of diagrams. More recently, physicists have reformulated pertubation theory calculations using spinor–helicity variables, in a method known as *on-shell recursion*. At tree-level, this method factorizes amplitudes involving n-many particles into products of scattering amplitudes with fewer than n-particles. At loop-level, on-shell recursion takes advantage of unitarity cuts to factorize loop amplitudes into lower order amplitudes. In this way, we arrive at general recursion relations for computing higher-order scattering processes.[28]

On-shell recursion illustrates how different choices of variables can make certain properties or patterns manifest. For instance, an elegant relationship known as the Parke–Taylor formula requires hundreds of pages to prove using Feynman diagrams but only a three-page inductive proof using the on-shell formulation. Progress in particle physics often comes from figuring out how to re-package perturbation series into ever more convenient forms, where otherwise-mysterious cancellations become clear. As noted in Section 6.1, some metaphysicians might be tempted to describe these examples as the result of finding a more fundamental language. In contrast, I agree with Woodward (2016, p. 1056) that metaphysical appeals to joint-carving do not give us a satisfying account of the relevant epistemological issues. The challenge is to understand how certain variable choices can make previously mysterious calculational patterns and cancellations intelligible.

[28]For background, see Henn and Plefka (2014), Dixon (2016), and Cheung (2017).

On-shell methods for scattering amplitudes illuminate a hidden symmetry that the tree-level superamplitudes possess in $\mathcal{N} = 4$ super Yang–Mills theory. This theory has an "obvious" superconformal symmetry $SU(2,2|4)$ that leaves its superamplitudes invariant.[29] Additionally, the tree-level superamplitudes of this theory possess a non-obvious *dual superconformal symmetry*. This hidden symmetry is also expressed by the symmetry group $SU(2,2|4)$, but now acting on a different set of variables defined in a different space than ordinary momentum variables (Elvang and Huang 2015, p. 95). Accounting for the intellectual significance of this hidden symmetry has numerous parallels to interpreting the hidden $SO(4)$ symmetry of the nonrelativistic hydrogen atom. As we saw in Section 6.8, an elementary presentation of the Hamiltonian for hydrogen does not make this hyperspherical symmetry manifest, although it does wear an "obvious" $SO(3)$ symmetry on the sleeves. Furthermore, this hidden symmetry is made manifest by moving to momentum variables. Similarly, in $\mathcal{N} = 4$ super Yang–Mills theory, the Lagrangian does not make manifest the hidden dual superconformal symmetry of the tree-level amplitudes. This hidden symmetry is made manifest by doing a series of variable changes, first moving to twistor space, then to a dual space, and finally to momentum twistor space. This section describes this series of variable transformations and the epistemic advantages we gain along the way.

In both examples, we can account for the intellectual significance of hidden symmetry in terms of epistemic dependence relations: moving to variables that make the symmetry manifest changes what it suffices to know to figure out if a given mathematical expression possesses the relevant symmetry. As discussed at the end of Section 6.4.1, by constructing objects that possess manifest dual superconformal symmetry, one can immediately infer that a more complicated object constructed from these invariant pieces also possesses this symmetry. There is surely part of this variable change that is merely convenient, but the change in epistemic dependence relations is also intellectually significant.

The main method for showing that superamplitudes possess a given symmetry is to show that the generators of that symmetry annihilate the superamplitudes. For instance,

[29] This symmetry group comprises a conformal part $SU(2,2)$ and an R-symmetry-part $SU(4)$. $SU(2,2)$ consists of 4×4 complex matrices of determinant one that preserve a Hermitian quadratic form of signature $(-1,-1,1,1)$. It is locally isomorphic to the conformal group $SO(2,4)$ of spacetime. The R-symmetry $SU(4)$ acts on the supersymmetry generators Q^A and $Q^\dagger_A$, where the index A ranges from one to four. These four 'supercharges' generate the supersymmetry transformations that transform bosons into fermions and vice versa.

tree-level superamplitudes possess Poincare symmetry because the ten generators of the Poincare group (four translations and six rotations/boosts) annihilate these amplitudes (Elvang and Huang 2015, p. 96). This holds for the super-Poincare group as well, which adds 16 fermionic supersymmetry generators Q^{Aa} and $\tilde{Q}_A^{\dot{a}}$. To show invariance under the superconformal group, the proof focuses on 16 additional fermionic conformal supersymmetry generators S_{Aa} and $\tilde{S}_{\dot{a}}^A$ along with properties of the momentum delta function and supermomentum Grassmann delta function (Elvang and Huang 2015, p. 99).

Changing to Twistor Variables

Representing the 62 superconformal symmetry generators of the graded Lie algebra $su(2,2|4)$ in spinor-helicity variables fails to treat these generators on equal footing. For instance, the translation generator has no derivative terms, the rotation/boost generators have one derivative term, and the conformal boost has two derivatives (Elvang and Huang 2015, p. 97). A desire to place these generators on equal footing with regards to derivative terms motivates the first change of variables. By moving to twistor variables, it is possible to provide a representation of these generators where every generator is a 1-derivative operator, which means that each has been linearized (Elvang and Huang 2015, p. 100). Just as the hydrogen atom case involves a Fourier transform from position space to momentum space, this variable change involves a Fourier transform from angle spinor variables to twistor variables. The resulting variables $\mathcal{W}_i^A$ are called supertwistors, and they consist of a triple of a square spinor, the Fourier transform of an angle spinor, and a Grassmann variable. This leads to a compact expression for every generator of the superconformal algebra where every generator is treated uniformly (Elvang and Huang 2015, p. 100). Furthermore, since the supertwistors scale homogenously under little group transformations, the resulting expression for the symmetry generators are invariant under this transformation.[30] This leads to a projective characterization of the twistors and supertwistors. The bosonic twistor part can be defined as a point in complex projective three space $\mathbb{CP}^3$. The supertwistors are points in $\mathbb{CP}^{3|4}$ space.

Changing variables to twistor space leads to a geometric interpretation of n-gluon tree-level amplitudes. It turns out that a tree-level gluon amplitude with q-many posi-

[30] In this context, the little group is the subgroup of the Poincare group that leaves the 4-momentum of a particle invariant. For massless particles, this is the two-dimensional Euclidean group $ISO(2)$, comprising translations in space and rotations around the direction of motion.

tive helicity gluons corresponds to a set of twistor points on a $(q-1)$-dimensional curve in bosonic twistor space $\mathbb{CP}^3$. For instance, anti-MHV amplitudes have two positive helicity gluons, so they correspond to a 1-dimensional curve. Hence, the amplitude itself comprises n twistors that lie on the same line in $\mathbb{CP}^3$ (Elvang and Huang 2015, p. 101).

Using Dirac's embedding formalism, one can provide an interpretation of twistors as a projective representation of spacetime points and null-lines. In the embedding formalism, the conformal group $SO(2,4)$ is realized as the Lorentz group of a six dimensional spacetime with metric $(-,-,+,+,+,+)$. Twistors are then defined as spinors on a conformal 4-dimensional subspace that satisfies a null condition $X \cdot X = 0$ and projectively identifies the 6-dimensional vector X with any scalar multiple rX. Each point X in this four-dimensional subspace is fixed by two twistor variables W_i and W_j. In other words, a line in twistor space corresponds to a point in the four-dimensional embedded spacetime. Conversely, any of two (six-dimensional) spacetime points X_i and X_j define a null-line, and they share the same twistor (since each twistor is identified with any scalar multiple of itself—resulting in twistor space being again $\mathbb{CP}^3$). Thus, a null-line in spacetime corresponds to a point in twistor space. In this way, twistor space is dual to the 4-dimensional embedded spacetime (dual in the same sense that lines and points are dual to each other in projective geometry).

Changing to Dual Coordinates

There are a few expressive disadvantages of twistor variables that motivate yet another variable change (taking us closer to making the hidden dual superconformal symmetry manifest). In the twistor variables, the translation generators of the Poincare group do not have a linear action on spinor variables. This means that the spinor variables are not invariant under translation, and hence both momentum and supermomentum are not automatically conserved in the supertwistor formalism. Instead, momentum and supermomentum conservation are enforced using delta functions (Elvang and Huang 2015, p. 103). Just as one of the motivations for spinor–helicity variables is to automatically enforce the on-shell condition (to "trivialize" this condition), the next variable change is motivated by a desire to automatically enforce conservation of momentum. We do this by interpreting momentum conservation geometrically, as a closed, convex contour, represented by a polygon. The momenta 4-vectors are directed edges of this n-sided polygon. The closure condition provides a geometric interpretation of the n-many momenta summing to zero.

A key step in characterizing the hidden symmetry is to move from the edges of this polygon (the momenta 4-vectors) to the dual notion, i.e. the points that define the vertices/cusps of the polygon. These dual coordinates y_i^μ define a dual space that although consisting of dual momentum variables is not itself characterized using spacetime coordinates. In this dual space, momentum conservation for scattering n-many particles is enforced by requiring that the $(n+1)$-th cusp y_{n+1} is the same as the first cusp of the polygon y_1, i.e. by requiring that the cusps are periodic (Elvang and Huang 2015, p. 103). Unlike the 4-momenta variables, these dual coordinates are invariant under translations. They thereby wear momentum conservation on their sleeves.

We proceed to re-express previous tree-level amplitude expressions using these dual space coordinates (and corresponding dual space coordinates for fermion variables). Since these amplitudes are now defined in dual space, it is possible to investigate a new class of symmetries, namely those encapsulated by dual superconformal symmetry. In this analysis, the dual inversion operator I plays a special role because the conformal boost generators $\mathcal{K}^\mu$ can be defined as intertwined with the translation operator by inversion: $\mathcal{K}^\mu = I\mathcal{P}^\mu I$. Since the dual coordinates are invariant under dual translation, this relationship makes it easy to see how the dual coordinates transform under other symmetry generators of the dual superconformal group. Ultimately, using the super-BCFW recursion relations re-expressed in these variables, it can be shown that all of the tree-level superamplitudes of $\mathcal{N} = 4$ SYM are invariant under dual superconformal symmetry. Even though this symmetry group is the same as that for regular superconformal symmetry, the symmetries are distinct. For instance, the tree amplitudes for gluon scattering in pure Yang–Mills theory are conformally invariant but not invariant under the corresponding symmetries of the dual conformal group (Elvang and Huang 2015, p. 105).

The two superconformal groups (ordinary and dual) can be combined into an even larger symmetry group known as the Yangian. This group has a countably infinite-dimensional algebra, where the lowest level generators correspond to the generators of the ordinary superconformal group. Since the tree-level superamplitudes are invariant under both ordinary and dual superconformal symmetry, they are ultimately invariant under the Yangian as well, manifesting an even larger hidden symmetry (Elvang and Huang 2015, p. 106).

Changing to Momentum Twistors

Despite demonstrating the dual superconformal symmetry by using dual coordinates, these coordinates are not ideal for expressing this symmetry. They do not themselves transform covariantly under the symmetry group (Elvang and Huang 2015, p. 107). Consequently, the resulting expressions for tree-level superamplitudes also do not wear this dual superconformal symmetry on their sleeves in dual coordinates. This is what ultimately motivates moving to *momentum twistors*.

Just as the twistor variables are geometrically dual to spacetime coordinates, the momentum twistors are geometrically dual to the dual coordinates y_i^{μ}. This means that a momentum twistor corresponds to a null-line in the dual y-space, and a point in the y-space corresponds to a line in the momentum twistor space. Furthermore, these momentum twistors transform as spinors, i.e. they have spinor indices. For convenience, we will call the momentum twistor space Z-space. The momentum twistor variables Z_i^I transform linearly under every transformation of the dual conformal group $SU(2,2)$, leading to a uniform and compact expression for the generators of this group (Elvang and Huang 2015, p. 108).

To re-express the amplitudes in a way that is manifestly invariant under dual conformal transformations, we form an invariant object out of the momentum twistor variables by contracting four of them with the Levi-Civita tensor for $SU(2,2)$. This leads to an invariant object called the 4-bracket, allowing us to re-express both the on-shell propagators and the tree-level amplitudes. Although the 4-bracket is convenient due to its symmetry properties, it is more than merely convenient: by building further objects (such as amplitudes) out of 4-brackets, it follows that these objects inherit the symmetry properties of dual conformal invariance. This is an instance of an epistemic dependence relation: to know that a resulting expression is invariant under the dual conformal group, it suffices to know that it is built out of component parts that are invariant.

By adding a corresponding Grassmann-variable to the momentum twistors, one forms momentum supertwistors, which make the dual superconformal symmetry manifest (Elvang and Huang 2015, p. 110). Here is a summary of the methodological upshot of all of these variable changes:

> Starting with the simple observation that momentum conservation is imposed in a rather ad hoc fashion, we introduced the auxiliary variables y_i such that momentum

> conservation is encoded in a geometric fashion. This led us to the realization of a new symmetry of the tree amplitude for $\mathcal{N}=4$ SYM, namely superconformal symmetry in the dual space y_i. The new symmetry set us on a journey to search for new variables, the momentum (super)twistors, that linearize the transformation rules. This culminated in the simple symmetric form of the n-point NMHV tree superamplitude. (Elvang and Huang 2015, p. 111)

Finally, the momentum twistors and momentum supertwistors have a further expressive property lacked by the dual coordinates in y-space. Although the y-space coordinates trivialize momentum conservation, they are nevertheless forced by hand to obey an algebraic constraint: $(y_i - y_{i+1})^2 = 0$. This enforces the on-shell momentum condition for the 4-momentum p_i. In contrast, the Z-coordinates are not subject to any analogous constraint. These coordinates are thereby defined freely in $\mathbb{CP}^3$. Working in this space of free Z-coordinates, we can study scattering amplitudes for n-many particles by picking *any* set of n-many points Z_i. To represent a scattering process, these points must ultimately form a closed contour, which means that each line (edge) is characterized by connecting subsequent points, i.e. (Z_i, Z_{i+1}). Due to the projectively dual relationship between y-space and Z-space, each of these lines (Z_i, Z_{i+1}) corresponds to a dual coordinate y_i. The fact that the contour is closed simply means that the nth line is (Z_1, Z_n), which entails the periodicity condition in dual coordinate space, i.e. that $y_{n+1} = y_1$. Recall that this periodicity condition simply means that momentum is conserved. In this way, our construction of a representation for scattering amplitudes in Z-space automatically enforces momentum conservation. Furthermore, the mapping of lines in Z-space to points in y-space forces adjacent y_i and y_{i+1} coordinates to obey an *incidence relation* that forces these adjacent y-coordinates to be null-separated. Since these adjacent coordinates are null-separated, the associated 4-momenta p_i are on-shell. In this way, the momentum twistor construction also automatically enforces that the represented scattering process is on-shell. This is a key difference with the y-space formalism itself, where the on-shell condition had to be enforced by hand (by requiring that adjacent y-coordinates be null-separated). Similar remarks apply for the momentum supertwistors (Elvang and Huang 2015, p. 112).

Moreover, momentum twistors provide a geometric interpretation of the propagators. In the dual space coordinates, propagators are expressed as $1/y_{ij}^2$, and a propagator is on-shell when $y_{ij}^2 = 0$. Using the aforementioned 4-bracket (which is a dual conformal invariant expressed in terms of four momentum twistors), the on-shell condition is re-expressed as requiring that the 4-bracket equals zero. Algebraically, this means that the four momentum twistors defining the 4-bracket are linearly dependent. Geometrically, this means that these four twistors belong to the same plane in $\mathbb{CP}^3$. This interpretation can be extended further by recasting propagator poles $y_{ij}^2 = 0$ as the intersections between certain lines and planes in momentum twistor space (Elvang and Huang 2015, p. 113).

To phrase this all more starkly: the reformulation using momentum twistors has enabled kinematic constraints (momentum conservation, on-shell momenta, and propagator poles) that were previously expressed *algebraically* (i.e. as solutions to equations) to be expressed *geometrically* (i.e. in terms of the intersections of lines and planes at certain points in momentum twistor space). This is yet another illustration of a difference in epistemic dependence relations. Rather than needing to know that a certain algebraic condition is satisfied by the variables of interest, the geometric reformulation shows that it suffices to know that a given geometric relationship holds. This is an instance of a much larger motif between algebraic and geometric expressive means that runs throughout various parts of mathematics. The interpretation of scattering amplitudes in terms of the volume of the amplituhedron takes this geometric reformulation even further. It shows that the equivalence of various representations of scattering amplitudes (derived from the BCFW recursion relations using different choices of line-shifts) is no coincidence, since they all correspond to different ways of triangulating a mathematical object known as the amplituhedron.

Insofar as physicists have an epistemic reason to trivialize certain conservation properties, they have an epistemic reason to privilege variables that do so. Imagine then that a physicist judges momentum twistor variables to be more fundamental than spinor–helicity variables. Rather than construing this judgment as involving a metaphysical commitment to joints in nature, we can apply the expressivist analysis from Section 6.6. In making this judgment of relative fundamentality, we implicitly endorse a set of norms on which one ought to prefer variables that can perform the various functional roles that momentum twistors perform (but that spinor–helicity variables cannot). The same could be said for viewing the dual coordinates as being more fundamental than spinor–helicity variables, since the dual coordinates make manifest the conservation of momentum.

6.10 Conclusion

We have seen that by changing variables, we can make otherwise obscured or hidden properties manifest. Against instrumentalism, I have argued that good variable choices can have non-instrumental epistemic value. Yet, the challenge of accounting for this epistemic value initially seems to favor fundamentalism. Here, I have shown that conceptualism has ample resources to accommodate the intellectual significance of making properties manifest. Good variable choices make intelligible properties of expressions and patterns in calculations. By changing variables, we sometimes make available new problem-solving plans, with concomitant differences in EDRs.

Sections 6.2-6.5 provided numerous elementary examples of making properties manifest. I showed how Cartesian coordinates make manifest the invariant degrees of freedom of horizontal and vertical lines. Likewise, polar coordinates make manifest properties of circles and diagonal lines. I provided a structurally similar illustration in the context of translating between natural languages.

Section 6.6 considered a rebuttal on behalf of fundamentalism. Scientists and mathematicians frequently judge one choice of variables to be more fundamental than another, especially in the context of making properties manifest. Hence, there is a burden on conceptualism to provide a non-metaphysical account of these practice-based judgments of fundamentality. Using expressivism, I provided one way to discharge this burden. To judge that a variable choice X is more fundamental than a variable choice Y is to express an attitude of *being for privileging* X over Y. If we focus on non-metaphysical reasons for privileging one variable choice over another, then this provides a non-metaphysically committal account of fundamentality.

Finally, Sections 6.8-6.9 developed two case studies of making a hidden symmetry manifest, concerning the hydrogen atom and supersymmetric Yang–Mills theory, respectively. In both cases, making the symmetry manifest requires transforming to new variables. Particularly in the case of supersymmetric Yang–Mills theory, we saw that one can use these new variables to construct objects that are manifestly invariant under the previously hidden symmetry. These manifestly-invariant objects can then be used to construct others, which inherit the property of being manifestly-invariant. Section 6.9 also illustrated numerous epistemic reasons that motivate physicists to transform variables, such as a desire to trivialize conservation of momentum.

Overall, conceptualism provides a promising account of the value of making symmetries manifest and of good variable choices generally. The success of conceptualism in this regard further undermines fundamentalism. Intuitively, manifest symmetries seem like a case where fundamentalism starts out with the upper hand. By showing that we can avoid metaphysically-committal notions of fundamentality *even in these cases*, we gain further reason to believe that we can avoid such commitments generally. If fundamentalism is not needed to account for the non-instrumental value of making properties manifest, it is hard to see where fundamentalism is needed—at least when it comes to assessing the value of compatible reformulations. Consequently, the arguments in this chapter insulate conceptualism from one of the strongest objections that a fundamentalist might leverage against it. Indeed, insofar as fundamentalists typically endorse Occam's razor, they should value the ontological parsimony of my conceptualist account.[31]

[31]Even if one views facts about metaphysical structure as part of a theory's ideology, rather than its ontology, Sider still advocates parsimony considerations here as well (2011, p. 14).

References

Abraham, R. and J. E. Marsden (1978). *Foundations of Mechanics.* 2nd ed. Benjamin/Cummings Publishing Company.

Abrahamsen, A. and W. Bechtel (2015). "Diagrams as Tools for Scientific Reasoning." *Review of Philosophy and Psychology* 6, pp. 117–131.

Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics.* Trans. by K. Vogtmann and A. Weinstein. 2nd ed. Springer.

Ash, A. and R. Gross (2008). *Fearless Symmetry: Exposing the Hidden Patterns of Numbers.* Princeton: Princeton University Press.

Avigad, J. (2015). "Mathematics and Language." In E. Davis and P. Davis (*eds.*), *Mathematics, Substance and Surmise: Views on the Meaning and Ontology of Mathematics.* Springer, pp. 235–256.

Baker, A. (2017). "Mathematics and Explanatory Generality." *Philosophia Mathematica* 25, pp. 194–209.

Balaban, A. (1985). "Applications of Graph Theory in Chemistry." *Journal of Chemical Information and Computer Sciences* 25, pp. 334–343.

Barker, S. (2012). "Expressivism about Making and Truth-making." In F. Correia and B. Schnieder (*eds.*), *Metaphysical Grounding. Understanding the Structure of Reality.* Cambridge: Cambridge University Press, pp. 272–293.

Barnes, E. (1992). "Explanatory Unification and the Problem of Asymmetry." *Philosophy of Science* 59, pp. 558–571.

Barrett, T. W. (2019). "Equivalent and Inequivalent Formulations of Classical Mechanics." *British Journal for Philosophy of Science* 70, pp. 1167–1199.

Batterman, R. W. (2002). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence.* Oxford: Oxford University Press.

Beddor, B. (2019). "Relativism and Expressivism." In M. Kusch (*ed.*), *The Routledge Handbook of Philosophy of Relativism.* Routledge, pp. 528–539.

Bethe, H. A. (1929). "Termaufspaltung in Kristallen." *Annalen der Physik* 3, pp. 133–208.

Bilalić, M., M. Graf, N. Vaci, and A. H. Danek (2019). "When the Solution is on the Doorstep: Better Solving Performance, but Diminished Aha! Experience for Chess Experts on the Mutilated Checkerboard Problem." *Cognitive Science* 43.

Black, M. (1970). "Induction and Experience." In L. Foster and J. W. Swanson (*eds.*), *Experience and Theory*. Amherst: University of Massachusetts Press, pp. 135–160.

Blackburn, S. (1993). *Essays in Quasi-Realism*. New York: Oxford University Press.

Blackburn, S. (1998). *Ruling Passions*. Oxford: Oxford University Press.

Bokulich, A. (2011). "How Scientific Models can Explain." *Synthese* 180, pp. 33–45.

Bokulich, A. (2018). "Representing and Explaining: The Eikonic Conception of Scientific Explanation." *Philosophy of Science* 85, pp. 793–805.

Bonolis, L. (2004). "From the Rise of the Group Concept to the Stormy Onset of Group Theory in the New Quantum Mechanics: A Saga of the Invariant Characterization of Physical Objects, Events and Theories." *Rivista del Nuovo Cimento* 27, pp. 1–110.

Brandom, R. B. (2001). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brandom, R. B. (2011). *Perspectives on Pragmatism: Classical, Recent, and Contemporary*. Cambridge, MA: Harvard University Press.

Brink, D. M. and G. R. Satchler (1968). "Graphical Methods in Angular Momentum." In *Angular Momentum*. 2nd ed. Oxford: Clarendon Press, pp. 112–136.

Burgess, C. and G. Moore (2006). *The Standard Model: A Primer*. Cambridge: Cambridge University Press.

Butterfield, J. N. (2006). "On Symmetry and Conserved Quantities in Classical Mechanics." In W. Demopoulos and I. Pitowsky (*eds.*), *Physical Theory and its Interpretation: Essays in Honor of Jeffrey Bub*. Dordrecht: Springer, pp. 43–99.

Carnap, R. (1935). *Philosophy and Logical Syntax*. London: Kegan Paul, Trench, Trubner.

Carnap, R. (1996 [1935]). *Philosophy and Logical Syntax*. Bristol: Thoemmes Press.

Cheung, C. (2017). "TASI Lectures on Scattering Amplitudes." *arXiv Preprint:1708.03872*.

Chrisman, M. (2007). "From Epistemic Contextualism to Epistemic Expressivism." *Philosophical Studies* 135, pp. 225–254.

Cohen, J. and C. Callender (2009). "A Better Best System Account of Lawhood." *Philosophical Studies* 145, pp. 1–34.

Cohen, S. (2016). "Theorizing about the Epistemic." *Inquiry* 59, pp. 839–857.

Colyvan, M. (2012). *An Introduction to the Philosophy of Mathematics.* Cambridge: Cambridge University Press.

Cornwell, J. F. (1984). *Group Theory in Physics.* Vol. 1. London: Academic Press.

Cotton, F. A. (1990). *Chemical Applications of Group Theory.* 3rd ed. New York: Wiley.

Craver, C. F. (2014). "The Ontic Account of Scientific Explanation." In M. Kaiser, O. Scholz, D. Plenge, and A. Hüttemann (*eds.*), *Explanation in the Special Sciences: The Case of Biology and History.* Dordrecht: Springer, pp. 27–52.

Curiel, E. (2014). "Classical Mechanics is Lagrangian; It is Not Hamiltonian." *British Journal for the Philosophy of Science* 65, pp. 269–321.

Dasgupta, S. (2018). "Realism and the Absence of Value." *Philosophical Review* 127, pp. 279–322.

Davidson, D. (1980 [1970]). "Mental Events." In *Essays on Actions and Events.* Oxford: Clarendon Press, pp. 207–227.

de Regt, H. (2009). "The Epistemic Value of Understanding." *Philosophy of Science* 76, pp. 585–597.

de Regt, H. (2017). *Understanding Scientific Understanding*. Oxford: Oxford University Press.

de Regt, H. and D. Dieks (2005). "A Contextual Approach to Scientific Understanding." *Synthese* 144, pp. 137–170.

de Regt, H., S. Leonelli, and K. Eigner (2009). "Focusing on Scientific Understanding." In *Scientific Understanding: Philosophical Perspectives.* Pittsburgh: University of Pittsburgh Press, pp. 1–17.

Detlefsen, M., D. K. Erlandson, J. C. Heston, and C. M. Young (June 1, 1976). "Computation with Roman Numerals." *Archive for History of Exact Sciences* 15, pp. 141–148.

Dixon, L. J. (2016). "A Brief Introduction to Modern Amplitude Methods." In *Journeys Through the Precision Frontier: Amplitudes for Colliders: TASI 2014 Proceedings of the 2014 Theoretical Advanced Study Institute in Elementary Particle Physics.* World Scientific, pp. 39–97.

Dunn, T. M., D. S. McClure, and R. G. Pearson (1965). *Some Aspects of Crystal Field Theory.* New York: Harper & Row.

Earman, J. and J. T. Roberts (2005). "Contact with the Nomic Part II: The Epistemological Argument for Humean Supervenience." *Philosophy and Phenomenological Research* 71, pp. 253–286.

Elgin, C. (2004). "True Enough." *Philosophical Issues* 14, pp. 113–131.

Elvang, H. and Y.-t. Huang (2015). *Scattering Amplitudes in Gauge Theory and Gravity.* Cambridge: Cambridge University Press.

Field, H. (2009). "Epistemology without Metaphysics." *Philosophical Studies* 143, pp. 249–290.

Field, H. (2018). "Epistemology from an Evaluativist Perspective." *Philosophers' Imprint* 18, pp. 1–23.

Figgis, B. N. and M. A. Hitchman (2000). *Ligand Field Theory and its Applications.* New York: Wiley-VCH.

Fock, V. A. (1935a). "Wasserstoffatom und nicht-Euklidische Geometrie." *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, pp. 169–188.

Fock, V. A. (1935b). "Zur Theorie des Wasserstoffatoms." *Zeitschrift für Physik* 98, pp. 145–154.

Fock, V. A. (2005). "On the Theory of the Hydrogen Atom." In *Linearity, Symmetry, and Prediction in the Hydrogen Atom.* Trans. by S. F. Singer. Springer, pp. 286–296.

Geroch, R. (1972). "Einstein Algebras." *Communications in Mathematical Physics* 26, pp. 271–75.

Gibbard, A. (1990). *Wise Choices, Apt Feelings.* Oxford: Oxford University Press.

Gibbard, A. (2003). *Thinking How to Live.* Cambridge, MA: Harvard University Press.

Gibbard, A. (2012). *Meaning and Normativity.* Oxford: Oxford University Press.

Gijsbers, V. (2007). "Why Unification is Neither Necessary nor Sufficient for Explanation." *Philosophy of Science* 74, pp. 481–500.

Gijsbers, V. (2013). "Understanding, Explanation, and Unification." *Studies in History and Philosophy of Science Part A* 44, pp. 516–522.

Goldfarb, W. (2003). *Deductive Logic.* 1st ed. Indianapolis: Hackett.

Goldman, A. I. (1986). *Epistemology and Cognition.* Cambridge, MA: Harvard University Press.

Goldman, A. I. (2017). "What can Psychology do for Epistemology?: Revisiting *Epistemology and Cognition*." *Philosophical Topics* 45, pp. 17–32.

Goldstein, H., C. P. Poole, and J. L. Safko (2002). *Classical Mechanics.* 3rd ed. San Francisco: Addison Wesley.

Graham, L. R. (2000). "Do Mathematical Equations Display Social Attributes?" *Mathematical Communities* 22, pp. 31–36.

Grimm, S. R. (2010). "The Goal of Explanation." *Studies in History and Philosophy of Science Part A* 41, pp. 337–344.

Grimm, S. R. (2014). "Understanding as Knowledge of Causes." In A. Fairweather (*ed.*), *Virtue Epistemology Naturalized. Bridges Between Virtue Epistemology and Philosophy of Science*. Dordrecht: Springer, pp. 329–345.

Halvorson, H. (2016). "Scientific Theories." In P. Humphreys (*ed.*), *The Oxford Handbook of Philosophy of Science*. Oxford: Oxford University Press, pp. 585–608.

Healey, R. (2007). *Gauging What's Real: The Conceptual Foundations of Contemporary Gauge Theories*. Oxford: Oxford University Press.

Hempel, C. G. (1965). "Aspects of Scientific Explanation." In *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: Free Press, pp. 331–496.

Hempel, C. G. and P. Oppenheim (1965 [1948]). "Studies in the Logic of Explanation." In *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: Free Press, pp. 245–290.

Henn, J. M. and J. C. Plefka (2014). *Scattering Amplitudes in Gauge Theories*. Vol. 883. Springer.

Hills, A. (2016). "Understanding Why." *Noûs* 50, pp. 661–688.

Hunt, J. (2014). "Group Theory or No Group Theory: Understanding Selection Rules in Atomic Spectroscopy." Bachelor's Thesis. University of Pittsburgh.

Hunt, J. (2016). "Indispensability and the Problem of Compatible Explanations." *Synthese* 193, pp. 451–467.

Hunt, J. (2021a). "Interpreting the Wigner–Eckart Theorem." *Studies in History and Philosophy of Science* 87, pp. 28–43.

Hunt, J. (2021b). "Understanding and Equivalent Reformulations." *Philosophy of Science* 88.

Hunt, J. (forthcoming). "Epistemic Dependence and Understanding: Reformulating Through Symmetry." *British Journal for the Philosophy of Science*.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.

Khalifa, K. (2012). "Inaugurating Understanding or Repackaging Explanation?" *Philosophy of Science* 79, pp. 15–37.

Khalifa, K. (2013). "The Role of Explanation in Understanding." *British Journal for the Philosophy of Science* 64, pp. 161–187.

Khalifa, K. (2015). "EMU Defended: Reply to Newman (2014)." *European Journal for Philosophy of Science* 5, pp. 377–385.

Khalifa, K. (2017). *Understanding, Explanation, and Scientific Knowledge.* Cambridge: Cambridge University Press.

Kitcher, P. (1989). "Explanatory Unification and the Causal Structure of the World." In P. Kitcher and W. Salmon (*eds.*), *Scientific Explanation.* Vol. 13. Minneapolis: University of Minnesota Press, pp. 410–505.

Korsgaard, C. (1983). "Two Distinctions in Goodness." *Philosophical Review* 92, pp. 169–195.

Kraut, R. (1990). "Varieties of Pragmatism." *Mind* 99, pp. 157–183.

Lancaster, T. and S. J. Blundell (2014). *Quantum Field Theory for the Gifted Amateur.* Oxford: Oxford University Press.

Lange, M. (2009a). *Laws and Lawmakers: Science, Metaphysics, and the Laws of Nature.* New York: Oxford University Press.

Lange, M. (2009b). "Why Proofs by Mathematical Induction are Generally not Explanatory." *Analysis* 69, pp. 203–211.

Lange, M. (2010). "What are Mathematical Coincidences (and Why Does it Matter)?" *Mind* 119, pp. 307–340.

Lange, M. (2013). "What Makes a Scientific Explanation Distinctively Mathematical?" *British Journal for the Philosophy of Science* 64, pp. 485–511.

Lange, M. (2014). "Aspects of Mathematical Explanation: Symmetry, Unity, and Salience." *Philosophical Review* 123, pp. 485–531.

Lange, M. (2017). *Because without Cause: Non-causal Explanations in Science and Mathematics.* New York: Oxford University Press.

Laudan, L. (1996). *Beyond Positivism and Relativism: Theory, Method, and Evidence.* Boulder, Colorado: Westview Press.

Lewis, D. (1983). "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61, pp. 343–377.

Lewis, D. (1984). "Putnam's Paradox." *Australasian Journal of Philosophy* 62, pp. 221–236.

Lewis, D. (1986). "Causal Explanation." In his *Philosophical Papers.* Vol. 2. New York: Oxford University Press, pp. 214–240.

Lipton, P. (2009). “Understanding without Explanation.” In H. de Regt, S. Leonelli, and K. Eigner (*eds.*), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press, pp. 43–63.

Loewer, B. (2007). “Laws and Natural Properties.” *Philosophical Topics* 35, pp. 313–328.

Lombrozo, T. (2011). “The Instrumental Value of Explanations.” *Philosophy Compass* 6, pp. 539–551.

Lutz, S. (2017). “What was the Syntax-Semantics Debate in the Philosophy of Science About?” *Philosophy and Phenomenological Research* 95, pp. 319–352.

MacFarlane, J. (2005). “The Assessment Sensitivity of Knowledge Attributions.” In T. S. Gendler and J. Hawthorne (*eds.*), *Oxford Studies in Epistemology*. Oxford: Oxford University Press, pp. 197–234.

Manders, K. (2008a). “Diagram-based Geometric Practice.” In P. Mancosu (*ed.*), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 65–79.

Manders, K. (2008b). “The Euclidean Diagram (1995).” In P. Mancosu (*ed.*), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 80–133.

Manders, K. (unpublished). “Expressive Means and Mathematical Understanding.”

Maudlin, T. (2018). “Ontological Clarity via Canonical Presentation: Electromagnetism and the Aharonov–Bohm Effect.” *Entropy* 20, pp. 465–485.

McIntosh, H. V. (1971). “Symmetry and Degeneracy.” In E. Loebl (*ed.*), *Group Theory and its Applications*. Vol. 2. New York: Academic Press, pp. 75–144.

Misak, C. (2013). *The American Pragmatists*. Oxford: Oxford University Press.

Morrison, M. (2000). *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge: Cambridge University Press.

Muller, F. A. (2005). “The Deep Black Sea: Observability and Modality Afloat.” *British Journal for the Philosophy of Science* 56, pp. 61–99.

Newman, M. P. (2013). “Refining the Inferential Model of Scientific Understanding.” *International Studies in the Philosophy of Science* 27, pp. 173–197.

Nolan, D. (1999). “Is Fertility Virtuous in its Own Right?” *British Journal for the Philosophy of Science* 50, pp. 265–282.

North, J. (2009). “The ‘Structure’ of Physics: A Case Study.” *The Journal of Philosophy* 106, pp. 57–88.

Pauli, W. (1926). “Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik.” *Zeitschrift für Physik* 36, pp. 336–363.

Potochnik, A. (2015a). "Causal Patterns and Adequate Explanations." *Philosophical Studies* 172, pp. 1163–1182.

Potochnik, A. (2015b). "The Diverse Aims of Science." *Studies in History and Philosophy of Science Part A* 53, pp. 71–80.

Potochnik, A. (2017). *Idealization and the Aims of Science*. Chicago: University of Chicago Press.

Price, H. (2011). *Naturalism without Mirrors*. Oxford: Oxford University Press.

Price, H. (2013). "The Descartes Lectures 2008." In *Expressivism, Pragmatism and Representationalism*. Cambridge: Cambridge University Press, pp. 3–64.

Putnam, H. (1979 [1965]). "Craig's Theorem." In *Mathematics, Matter and Method. Philosophical Papers*. 2nd ed. Vol. 1. Cambridge: Cambridge University Press, pp. 228–236.

Putnam, H. (1979 [1971]). "Philosophy of Logic." In *Mathematics, Matter and Method. Philosophical Papers*. 2nd ed. Vol. 1. Cambridge: Cambridge University Press, pp. 323–358.

Putnam, H. (1979 [1968]). "The Logic of Quantum Mechanics." In *Mathematics, Matter and Method. Philosophical Papers*. 2nd ed. Vol. 1. Cambridge: Cambridge University Press, pp. 174–197.

Railton, P. (1981). "Probability, Explanation, and Information." *Synthese* 48, pp. 233–256.

Rosen, G. (1994). "What is Constructive Empiricism?" *Philosophical Studies* 74, pp. 143–178.

Rosenstock, S., T. W. Barrett, and J. O. Weatherall (2015). "On Einstein Algebras and Relativistic Spacetimes." *Studies in History and Philosophy of Modern Physics* 52, pp. 309–16.

Rosenstock, S. and J. O. Weatherall (2016). "A Categorical Equivalence between Generalized Holonomy Maps on a Connected Manifold and Principal Connections on Bundles over that Manifold." *Journal of Mathematical Physics* 57.

Salmon, W. C. (1989). *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.

Salmon, W. C. (1998). "Scientific Explanation: Causation and Unification." In *Causality and Explanation*. Oxford: Oxford University Press, pp. 68–78.

Salmon, W. C. (1998 [1984]). "Scientific Explanation: Three Basic Conceptions." In *Causality and Explanation*. Oxford: Oxford University Press, pp. 320–329.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

Schlimm, D. and H. Neth (2008). “Modeling Ancient and Modern Arithmetic Practices: Addition and Multiplication with Arabic and Roman Numerals,” *CogSci*, pp. 2097–2102.

Scholz, E. (2006). “Introducing Groups into Quantum Theory (1926–1930).” *Historia Mathematica* 33, pp. 440–490.

Schroeder, M. (2008). *Being For.* Oxford: Oxford University Press.

Searle, J. R. (1995). *The Construction of Social Reality.* New York: Free Press.

Sellars, W. (1958). “Counterfactuals, Dispositions, and the Causal Modalities.” In H. Feigl, M. Scriven, and G. Maxwell (*eds.*), *Concepts, Theories, and the Mind-Body Problem.* Vol. 2. Minneapolis: University of Minnesota Press, pp. 225–308.

Sethna, J. P. (2006). *Statistical Mechanics: Entropy, Order Parameters, and Complexity.* Oxford: Oxford University Press.

Sider, T. (2011). *Writing the Book of the World.* Oxford: Oxford University Press.

Siegel, W. (2005). *Fields.* 3rd ed.

Singer, S. F. (2005). *Linearity, Symmetry, and Prediction in the Hydrogen Atom.* Springer.

Skow, B. (2016). *Reasons Why.* Oxford: Oxford University Press.

Sosa, E. (2015). *Judgment and Agency.* Oxford: Oxford University Press.

Srednicki, M. (2007). *Quantum Field Theory.* Cambridge: Cambridge University Press.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation.* Cambridge, MA: Harvard University Press.

Strevens, M. (2013). “No Understanding without Explanation.” *Studies in History and Philosophy of Science Part A* 44, pp. 510–515.

Strocchi, F. (2013). *An Introduction to Non-Perturbative Foundations of Quantum Field Theory.* Oxford: Oxford University Press.

Styer, D. F. et al. (2002). “Nine Formulations of Quantum Mechanics.” *American Journal of Physics* 70, pp. 288–297.

Tappenden, J. (2008). “Mathematical Concepts: Fruitfulness and Naturalness.” In P. Mancosu (*ed.*), *The Philosophy of Mathematical Practice.* Oxford: Oxford University Press, pp. 276–301.

Thomasson, A. L. (2020). *Norms and Necessity.* Oxford: Oxford University Press.

Torre, C. G. (1997). “Local Cohomology in Field Theory: With Applications to the Einstein Equations.” *arXiv Preprint: hep-th/9706092.*

Trinajstić, N. (1992). *Chemical Graph Theory*. 2nd ed. Boca Raton: CRC Press.

Trout, J. D. (2007). "The Psychology of Scientific Explanation." *Philosophy Compass* 2, pp. 564–591.

van Fraassen, B. C. (1975). "Platonism's Pyrrhic Victory." In A. R. Anderson, R. B. Marcus, and R. M. Martin (*eds.*), *The Logical Enterprise*. New Haven: Yale, pp. 39–50.

van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Oxford University Press.

van Fraassen, B. C. (1994). "Gideon Rosen on Constructive Empiricism." *Philosophical Studies* 74, pp. 179–192.

van Fraassen, B. C. (1995). "'World' is Not a Count Noun." *Noûs* 29, pp. 139–157.

Weatherall, J. O. (2016). "Are Newtonian Gravitation and Geometrized Newtonian Gravitation Theoretically Equivalent?" *Erkenntnis* 81, pp. 1073–91.

Weatherall, J. O. (2019). "Theoretical Equivalence in Physics: Part 1." *Philosophy Compass* 14.

Wedgwood, R. (2002). "The Aim of Belief." *Philosophical Perspectives* 16, pp. 267–297.

Wigner, E. P. (1931). *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*. Braunschweig: F. Vieweg & Sohn.

Wigner, E. P. (1959). *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. Trans. by J. J. Griffin. New York: Academic Press.

Wittgenstein, L. (2009 [1953]). *Philosophical Investigations*. Trans. by G. Anscombe, P. Hacker, and J. Schulte. 4th ed. Wiley-Blackwell.

Wittgenstein, L. (2009 [1949]). *Philosophy of Psychology: A Fragment*. Trans. by G. Anscombe, P. Hacker, and J. Schulte. 4th ed. Wiley-Blackwell.

Wittgenstein, L. (1993 [1931]). "Remarks on Frazer's *Golden Bough*." In J. C. Klagge and N. Alfred (*eds.*), *Philosophical Occasions. 1912–1951*. Indianapolis: Hackett.

Woodard, E. (forthcoming). "Why Double Check?" *Episteme*.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. (2016). "The Problem of Variable Choice." *Synthese* 193, pp. 1047–1072.

Woodward, J. and C. Hitchcock (2003a). "Explanatory Generalizations, Part I: A Counterfactual Account." *Noûs* 37, pp. 1–24.

Woodward, J. and C. Hitchcock (2003b). "Explanatory Generalizations, Part II: Plumbing Explanatory Depth." *Noûs* 37, pp. 181–199.

Yamagishi, K. (1997). "When a 12.86% Mortality is More Dangerous than 24.14%: Implications for Risk Communication." *Applied Cognitive Psychology* 11, pp. 495–506.