OXFORD

the Practical
Origins *of* Ideas

genealogy as conceptual reverse-engineering

MATTHIEU QUELOZ

The Practical Origins of Ideas

The Practical Origins of Ideas

Genealogy as Conceptual Reverse-Engineering

MATTHIEU QUELOZ

For my parents

Contents

Acknowledgements

List of Abbreviations

AC   *The Anti-Christ*. Friedrich Nietzsche. 2005. Translated by Judith Norman. Edited by Aaron Ridley and Judith Norman. Cambridge: Cambridge University Press.

BGE   *Beyond Good and Evil*. Friedrich Nietzsche. 2002. Translated by Judith Norman. Edited by Rolf-Peter Horstmann and Judith Norman. Cambridge: Cambridge University Press.

BT   *The Birth of Tragedy*. Friedrich Nietzsche. 1999. Translated by Ronald Speirs. Edited by Raymond Geuss and Ronald Speirs. Cambridge: Cambridge University Press.

D   *Daybreak: Thoughts on the Prejudices of Morality*. Friedrich Nietzsche. 1997. Translated by R. J. Hollingdale. Edited by Maudemarie Clark and Brian Leiter. Cambridge: Cambridge University Press.

EH   *Ecce Homo*. Friedrich Nietzsche. 2005. Translated by Judith Norman. Edited by Aaron Ridley and Judith Norman. Cambridge: Cambridge University Press.

eKGWB   *Digital Critical Edition of the Complete Works and Letters*. Friedrich Nietzsche. 2009. Edited by Paolo D'Iorio. Based on the critical text by G. Colli and M. Montinari. Berlin: De Gruyter. www.nietzschesource.org/eKGWB.

EPM   *An Enquiry Concerning the Principles of Morals*. David Hume. 1998. Edited by Tom L. Beauchamp. Oxford: Clarendon Press.

GM   *On the Genealogy of Morality*. Friedrich Nietzsche. 1998. Translated by Maudemarie Clark and Alan J. Swensen. Indianapolis, IN: Hackett.

GS   *The Gay Science*. Friedrich Nietzsche. 2001. Translated by Josefine Nauckhoff and Adrian Del Caro. Edited by Bernard Williams. Cambridge: Cambridge University Press.

HA   *Human, All Too Human*. Friedrich Nietzsche. 1986. Translated by R. J. Hollingdale. Edited by Richard Schacht. Cambridge: Cambridge University Press.

T   *A Treatise of Human Nature*. David Hume. 2000. Edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.

TI   *Twilight of the Idols*. Friedrich Nietzsche. 2005. Translated by Judith Norman. Edited by Aaron Ridley and Judith Norman. Cambridge: Cambridge University Press.

TL   'On Truth and Lies in a Nonmoral Sense'. Friedrich Nietzsche. 1979. In *Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870's*. Edited and translated by Daniel Breazeale, 79–97. Atlantic Highlands, NJ: Humanities Press International.

UM   *Untimely Meditations*. Friedrich Nietzsche. 1997. Edited and translated by
          Daniel Breazeale. Cambridge: Cambridge University Press.
WS   *The Wanderer and His Shadow*. Friedrich Nietzsche. 1986. Translated by Judith
          Norman. Edited by Richard Schacht. Cambridge: Cambridge University Press.

Our common stock of words embodies all the distinctions men have found worth drawing, and the connections they have found worth making, in the lifetimes of many generations.

J. L. Austin, 'A Plea for Excuses', 130

Concepts are neither true nor false, but they can be evaluated: do we have reason to track the distinction drawn by the concept? Should we have this or that concept in our repertoire at all?

S. Haslanger, 'Social Meaning and Philosophical Method', 28

1

Why We Came to Think as We Do

We did not make the ideas we live by. They are, for the most part, ideas we inherited, unthinkingly growing into patterns of thought cultivated by others, with little sense of why just these ways of seeing, valuing, and reasoning should have gained hold in the first place. Some ideas, like that of *water*, may be so plainly useful for creatures like us as to appear inevitable. But many of our most venerable ideas—such as *truth*, *knowledge*, or *justice*—are highly abstract, and their practical value for us is elusive. Why did these 'highest concepts', these 'last wisps of smoke at the evaporating end of reality' (*TI*, Reason, §4), as Nietzsche called them, ever become so important to us? What was the point of coming to think in terms of these grand abstractions, and what would we lose if we lacked them?

Such *Pragmatic Questions* about the practical origins of ideas have seldom been raised. They have tended to be side-lined by more traditional *Socratic Questions* of the form 'What is X?' Aiming straight at the essence of truth, knowledge, or justice, the Socratic approach reckons that if only we achieve clarity about what these things really are, an understanding of why we came to be concerned with them will follow. Socratic Questions can prove obstinately vexing, however, and a consensus on what truth, knowledge, or justice are has yet to emerge. Accordingly, some have concluded with the American pragmatist C. S. Peirce that 'we must not begin by talking of pure ideas—vagabond thoughts that tramp the public high-ways without any human habitation—but must begin with men and their con-versation' (1931, 8.112). Peirce, like the philosophers I discuss in this book, diagnosed a tendency in philosophy to set ideas too high above human affairs, to contemplate them entirely *in vacuo*. Ideas are in their element in distinctive contexts of purposive human action, action that takes place against a background of contingent facts about us and the world we live in. Trying to understand the ideas we live by in isolation from the circumstances in which they are felicitously deployed is like studying a shoal of beached fish as if they were in their natural habitat.

Instead, we can turn the order of explanation around and let the *what* grow out of the *why*: we approach the question of the nature of truth, knowledge, or justice by first asking why we came to think in these terms. Such an inquiry into the origins of ideas can take many guises. Plato asked after the origins of ideas, but he sought them in an abstract realm of Forms. Conceptual historians of various stripes asked after the origins of ideas, but they sought them by tracing the changing

meanings of words across different socio-historical contexts.[1] My concern, by contrast, is with the *practical origins* of ideas: with the ways in which the ideas we live by can be shown to be rooted in practical needs and concerns generated by certain facts about us and our situation.

If an idea persists, the reason may be that it fills a need, or at least that it earns its keep through subservience to *some* kind of concern or interest. What motivates this line of inquiry is the realization that we are, as Jane Heal puts it, 'finite in our cognitive resources while the world is immensely rich in kinds of feature and hence in the possibilities it offers for conceptualization' (2013, 342). Why do we find at our disposal just the concepts we do rather than any of the countless imaginable alternatives? As Heal goes on to remark, this question cannot be answered simply by observing that using certain concepts enables us to form *true judgements* in terms of those concepts. More needs to be said—in particular, about what makes thinking and judging in just these terms *worthwhile*. This is especially true of the abstract notions at the heart of philosophy, which seem to be the stuff of idle grandiloquence rather than effective action. What needs, if any, were filled by introducing these ideas into our repertoire? What necessity was the mother of these inventions?

The method I propose to explore in this book is designed to help us look at ideas from a practical point of view—to look at what ideas *do* rather than at whether the judgements they figure in are *true*—in order to see how exactly our ideas are bound up with our needs and concerns. This method, which I propose to call *pragmatic genealogy*, consists in telling partly fictional, partly historical narratives exploring what might have driven us to develop certain ideas in order to discover what these ideas do for us. What *point* do they serve? What is the *salient useful difference* these ideas make to the lives of those who live by them?[2] Much as

---

[1] An important early example of a historicizing approach to philosophical concepts is Gustav Teichmüller's *Studien zur Geschichte der Begriffe* (1864). More recent examples include the *Historisches Wörterbuch der Philosophie* (1971–2007) and the works in the *Oxford Philosophical Concepts* series. For histories of concepts in social and political thought, see *Geschichtliche Grundbegriffe* (1972–97) as well as the genealogies of concepts presented by Skinner (2002, 2009) and other representatives of the Cambridge School; for scientific thought, the seminal work of Gaston Bachelard and Georges Canguilhem lies upstream of numerous forms of conceptual history, including historical epistemology and the history and philosophy of science (HPS). Underneath the umbrella term 'conceptual history' thus reside substantially different approaches that I cannot explore further here. Another relevant tradition whose relation to my topic merits more attention than I can give it is the sociology of knowledge, especially as exemplified by Karl Mannheim's *Ideology and Utopia* (1936).

[2] In what follows, usefulness will primarily be cashed out in terms of the tendency to satisfy individual and social needs, but pragmatic genealogies could be told in relation to all kinds of practical concerns and interests. My focus on needs is not meant to exclude the possibility of using pragmatic genealogy to criticize conceptual practices by revealing their pernicious subservience to problematic interests; nor does the use of a state-of-nature fiction commit one to telling a vindicatory story—think of Rousseau's (1977) genealogy of inequality as interpreted by Neuhouser (2014). But since the primary purpose of the pragmatic genealogists discussed in this book is to vindicate and defend practices, and since, for that purpose, it makes sense for them to focus on genuine needs (see Chapter 9), I shall follow them and concentrate on needs. As we shall see, moreover, even a genealogy rooted in needs can have revisionary implications.

an archaeologist who digs up a mysterious relic will try to reverse-engineer its point by imaginatively reconstructing the life of those who used it and hypothesizing what useful difference it might have made to that life, we can take an abstract idea whose point eludes us, such as *truth*, *knowledge*, or *justice*, and try to explain why we came to think in these terms by reconstructing the practical problems that these ideas offer practical solutions to. A pragmatic genealogy answers the question of why we came to think as we do by reverse-engineering the points of ideas, tracing them to their practical origins, and revealing what they do for us when they function well.

In asking why we came to think as we do, we are then not concerned with what triggered the occurrence of an idea, why an individual applied the idea as they did, or why they came to acquire the idea in a community in which it was already common currency.³ Answers to any of these questions would themselves draw on the idea at issue and thereby presuppose what we are trying to explain, namely the *communal acceptance* of the idea, the *practice* of thinking in these terms in the first place.⁴ Thus, the explananda whose practical origins we are investigating are the *communal practices of living by* certain ideas, such as the practice of living by the concept of knowledge, or the value of truth, or the virtue of justice. To *live by* a concept, value, or virtue is not just to understand it—that can be done in a *disengaged* way, as an ethnographer understands an idea at work in a different culture by imaginatively inhabiting that culture's viewpoint without making it her own. One can understand the idea behind some religious festivity without living by it oneself. To live by an idea is to be an *engaged* user of it, where that entails being responsive, in the conduct of one's own affairs, to the reasons generated by the idea's applicability in given situations.⁵

Let the term 'conceptual practice' therefore stand for a community's practice of letting its thoughts, attitudes, and actions be shaped and guided by a given idea. Unlike mere practices, such as walking on one's feet rather than one's hands, *conceptual* practices are essentially shaped by sensitivity to conceptual norms or reasons—take away the idea in terms of which those norms and reasons are articulated, and the practice collapses. The term's emphasis on the conceptual is apt even if it is taken to cover values and virtues, for valuing things involves value concepts, and while possessing a virtue may be a matter of reliably manifesting a

---

³ This would lead into a rather different set of debates over the origins of concepts, namely those focusing on whether concepts are learned or innate. See Carey (2009).

⁴ There is a distant echo here of the Wittgensteinian distinction between justifying a thought or an action by reference to the criteria or rules encoded in a concept, which is often straightforward, and the more vexing task of justifying the practice of operating according to such a concept. For a discussion of Wittgenstein's views on this, see Queloz (2016).

⁵ The engaged/disengaged terminology and the notion of 'living by' an idea hail from Moore (2006), who develops them in his discussion of thick concepts on the basis of Williams's thoughts on the 'ethnographic stance' (1986, 203–4; 1995d, 207; 2011, 157). The engaged/disengaged distinction admits of degrees, and allows for pretence, role-playing, and ironic uses: see Moore (2006, 138).

disposition out of motives that make no reference to that virtue (the genuinely humble person is precisely not motivated by a desire to appear humble), its systematic recognition and cultivation *as* a virtue by a community requires the concept *of* that virtue, which is why even virtues come as complexes of dispositions and concepts that are helpfully bundled under the heading of 'conceptual practices'.[6]

Once we understand inquiry into the origins of ideas as inquiry into the practical origins of conceptual practices, we can ask of any such practice why it arose and what it does for those who engage in it. This is the spirit that led Voltaire to the conclusion that if God did not exist, we would have to invent him.[7] It is a *pragmatic* spirit, because it focuses primarily on human *practices* of living by certain ideas rather than on what these ideas refer to, and because it seeks to make sense of those practices in terms of their *practical point*. The Socratic Question then cedes priority to the Pragmatic Question: 'Why would creatures like us be driven to develop the idea of X?' In Voltairean terms: if it did not exist, why would we have to invent it?

1.1 Bringing the Pragmatic Genealogical Tradition into View

One of my two main objectives in this book is to uncover a methodological tradition that pursues this Pragmatic Question by telling pragmatic genealogies of conceptual practices. Perhaps surprisingly, this tradition cuts across the analytic–continental divide and runs right through the heartland of Anglophone philosophy. The reason this may come as a surprise is that there is a perceived rift in philosophy between 'those who think that everything must be genealogised' and 'those who think that there is nothing to be learned from genealogy' (Srinivasan 2015, 326). The rift is often said to line up with the analytic–continental divide, the implication being that genealogy is something 'continental', and that Anglophone philosophy—certainly in the analytic tradition—defines itself through its opposition to genealogy.[8]

---

[6] What exactly concepts, values, and virtues should be taken to be in this context will become clearer in later chapters. On the need for interdependent concepts, dispositions, and practices to co-evolve, see Pettit (2018, 25–8).

[7] See Voltaire (1877–85, Vol. 10, 402).

[8] See Blackburn (2005), Boghossian (2006), Dutilh Novaes (2015), Glock (2008a, b), and Srinivasan (2011, 2). The divide itself has long been recognized as a strange cross-classification of the methodological and the topographical, as well grounded as that between cars with a four-wheel drive and cars from Japan (Williams 2006b, 201); moreover, the origins of analytic philosophy anyway lie on the continent (Dummett 1993; Glock 2008b). Yet there is some truth to the idea that an ahistorical spirit was characteristic of analytic philosophy in its early days. Wittgenstein emphatically captured it when he wrote in 1916: 'What has history to do with me? Mine is the first and only world!' (1979, 82). For a discussion of Wittgenstein's attitude towards history, see Glock (2006b, 2008a, b, 2017a) and Sluga

Subverting this divide, I want to suggest that Anglophone philosophy has its own genealogical tradition to look back on. This tradition of pragmatic genealogy remains invisible as long as we conceive of genealogy either along Foucauldian lines as something approaching or even exemplifying regular historiography, as Alexander Nehamas does when he suggests that Nietzschean 'genealogy simply *is* history, correctly practised' (1985, 246n1),[9] or along Hobbesian lines as a purely justificatory contrast foil that serves to exhibit the present as preferable to some hypothetical state of nature. A more helpful entry-point for our purposes is offered by the *Oxford Dictionary of Philosophy*'s definition of 'genealogy':

> Genealogy is part historical reconstruction of the way certain concepts have come to have the shape they do, and part 'rational reconstruction' or story about the function they serve, which may or may not correspond to historical evolution.   (Blackburn 2016)[10]

This broader characterization of genealogy as containing two aspects that can be combined in various proportions allows us to situate genealogical explanations on a spectrum rather than on either side of a dichotomy. At one end of the spectrum, we can place the fictional and primarily justificatory state-of-nature stories in political philosophy from Hobbes (2006) through Locke (2003) to Nozick (1974): these are genealogical insofar as they are developmental narratives, but it is for the most part doubtful that these narratives have any serious explanatory ambitions.[11] They are more plausibly read as justificatory arguments in genealogical guise. At the other end of the spectrum, we have the thoroughly historical and primarily explanatory genealogies of a Foucauldian stripe which are particularly popular outside philosophy: these do not start out from a state of nature, but trace out the multiple roots of something across real history; and while such genealogies tend to

---

(1998). Glock also compares Wittgenstein's 'remarks on the natural history of human beings' (2009, §415) with Williams's genealogical method (2006b, 296–303).

[9] Something approaching this conception of genealogy as regular history is endorsed by Geuss (1999, 22–3), Owen (2007, 143), Sarasin (2008), and Migotti (2016). Of course, if one's notion of history is broad enough, it is trivially true that genealogy is history; but this then precisely masks the difference—which is surely there, and which the aim of understanding what is specific to genealogy requires us to bring out—between history as practised by academic historians and history as practised by philosophical genealogists.

[10] For a fuller characterization of genealogy that serves my purposes equally well, see Owen (2010). For other characterizations of genealogy that situate it closer to history and further from functional explanation, see Bevir (2008) and Saar (2002, 2007, 2008).

[11] See Kavka (1986), Hampton (1987), and Angehrn (2007). Following Kant's interpretation of Hobbes, the point of hypothetical state-of-nature stories has often been thought to lie in demonstrating the *exeundum* principle—the rational imperative of exiting the *status naturalis* to live under the law (Byrd and Hruschka 2010; Kersting 2017, 30; Mori 2017, 104). It is true that Nozick (1974) explicitly harbours explanatory ambitions as well as justificatory ones. But see Williams (2002, ch. 2) for a critical discussion of Nozick's explanatory ambitions.

strip off the veneer of inevitability by revealing the contingency of present arrangements, this is usually offered as factual input to normative reflection, presenting genealogy as merely preparatory rather than constitutive of critique.[12]

But halfway between these two poles, we find genealogical explanations that combine fiction with history and justification with explanation. These hybrid genealogies fall squarely into the core territory of Anglophone philosophy, and notably include David Hume's genealogy of the virtue of justice (*T*), Edward John Craig's genealogy of the concept of knowledge (1990, 1993), Bernard Williams's genealogies of the virtues of truth (2002) and of the political idea of liberty (2005c), and Miranda Fricker's genealogy of the virtue of testimonial justice (2007). Following this trajectory, we could add Michael Hannon's (2013, 2015, 2019) and Steven Reynolds's (2017) Craigean genealogies of the concepts of knowledge and understanding, Maria-Sibylla Lotter's genealogy of the concept of the person (2012), Thomas Simpson's genealogy of the concept of trust (2012), Martin Kusch and Robin McKenna's genealogy of relativism and absolutism (2018), Philip Pettit's genealogy of moral desirability and responsibility (2018), and no doubt many others.[13] What these genealogists have in common is that they start from some fictional state of nature (or some equivalent of it); they seek to explain our ways of going on by presenting them as elaborations of prototypes that, in such state-of-nature situations, creatures like us would be driven to develop in virtue of certain needs; and they draw some normative guidance from these insights, because these put them in a position to assess whether we now share and endorse these needs, and whether the ideas are worth continued cultivation.

Once this line of continuity is rendered salient, it is easy to see much the same kind of genealogical method at work in the early writings of Friedrich Nietzsche, with whom the philosophical use of genealogy remains indelibly associated. I shall argue that in his Basel period (1869–79), Nietzsche sketches pragmatic genealogies that similarly start out in something like a fictional state of nature, and aim to make sense of ideas of justice or truthfulness in terms of their practical value for creatures like us. This leads on to the question whether Nietzsche's later genealogical method refines rather than replaces this earlier use of the

---

[12] See, e.g., Dutilh Novaes (2015, 100–1). While stopping short of yielding normative conclusions, even such Foucauldian genealogies do not leave everything as it was. Koopman (2013, 95) argues that it is not so much the fact of contingency as the way in which something contingently arose which is of interest, because it makes explicit and opens up to critique the enabling background assumptions of practices (2013, 21). Others have emphasized the ways in which genealogies uncover the role of social power in the construction of the self (Saar 2007), render practices uncanny (Menge 2017), or 'possibilize' them by enabling us to isolate, from the contingent processes that led us where we are, the possibility of no longer thinking as we do (Lorenzini forthcoming).

[13] The genealogies of ethics proposed by Joyce (2006), Prinz (2007), Kitcher (2011), Tomasello (2016), and Brandhorst (2021) are also broadly pragmatist and naturalistic in spirit, but they are in the business of forming historical conjectures as to how (elements of) ethics might in fact have arisen—they are primarily about our hominin past, whereas the pragmatic genealogies I am concerned with treat this as being at most a secondary application of models whose primary purpose is to elucidate our present.

method. But what, in any case, these early genealogies show us is that there is a pragmatic genealogical tradition which cuts across the analytic–continental divide.

Speaking of a 'tradition' raises the question whether there really was, as the etymology of the word suggests, a *passing on* of ideas from X to Y in the sense of a fairly direct relation of influence—though the term is also used more loosely, for example when it is said that, unbeknownst to Y, Y produced something *in the tradition of* X. Many an amateur poet composes verse in the tradition of Petrarch without knowing it (an example that also illustrates the notorious difficulty of ascertaining influence relations).[14] In the case of the tradition that forms the topic of this book, it would seem to me to be worth speaking of a pragmatic genealogical tradition even if there were no direct lines of influence between its members, because it would serve to highlight that besides what people widely regard as *the* tradition of using a genealogical method in philosophy (the tradition notably involving Foucault and the later Nietzsche as seen through a Foucauldian lens), there is *another* genealogical tradition, if only in the sense that there is a series of philosophers who all use a different genealogical method in strikingly similar ways.

As it happens, however, there are fairly direct relations of influence between the genealogists I discuss. Miranda Fricker explicitly develops Williamsian and Craigean ideas in her *Epistemic Injustice*, and her dissertation at Oxford in the 1990s was jointly supervised by Sabina Lovibond and Bernard Williams. She recalls Williams handing her his copy of Craig's *Knowledge and the State of Nature* and roundly recommending it.[15] Williams would also begin lectures on unrelated topics by recommending Craig's book, which he acknowledges as an inspiration for his *Truth and Truthfulness* (the book's subtitle, *An Essay in Genealogy*, echoes Craig's subtitle, *An Essay in Conceptual Synthesis*).[16] Williams and Craig were colleagues at Cambridge from 1967 to 1988 and mutually influenced each other. Craig explicitly draws on the 1973 paper in which Williams argues that the natural home for the concept of knowledge is not the situation of the *examiner* who assesses whether someone knows something already known to the examiner, but rather that of the *inquirer* who seeks to identify someone who knows something the inquirer does not yet know. Williams also sketched a precursor to Craig's genealogy of the concept of knowledge in his 1978 book on Descartes.[17]

---

[14] For an account that sets the bar for influence relations particularly high, see Skinner (1966). For a critique of Skinner's account as being sceptical to the point of being disabling, see Oakley (1999, 138–87).

[15] Fricker, personal communication.

[16] See Williams (2002, 21, 32–3). For the claim that Williams would start his lectures by recommending Craig's book, see Millgram (2009, 162n21).

[17] See Craig (1990, 18), which references Williams (1973a, 146). For a precursor to Craig's genealogy of the concept of knowledge, see Williams (2005b, ch. 2). Williams later wrote that he got the idea for a pragmatic genealogy of the concept of knowledge 'from the Australian philosopher Dan Taylor, who

Going further back, Williams was a Nietzsche scholar and Craig a Hume scholar, so that there can be no doubt that they were serious readers of Nietzsche and Hume, respectively. As for Nietzsche's relation to Hume, it is harder to make out, refracted as it may have been through Hume's influence on Darwin and Darwin's influence on Nietzsche's close friend Paul Rée.[18] But as we shall see in Chapter 5, Nietzsche presents himself as improving on the method of the 'English genealogists' that he read about in his extensively annotated copy of W. E. H. Lecky's *History of European Morals* (1869), a book that notably outlines the views of Hume. Nietzsche also read some Hume first-hand (though never, as far as we know, the *Treatise*), and first encountered Hume's ideas in the summer of 1865, when he was a student in Bonn and attended Carl Schaarschmidt's lectures (among the recommended readings was Albert Schwegler's *Geschichte der Philosophie im Umriß* (1848), which includes a chapter on Hume). He also carefully studied Maximilian Drossbach's Humean account of causality as well as the histories of philosophy by Kuno Fischer, Friedrich Ueberweg, and F. A. Lange, in which Hume featured prominently, if only as a stepping-stone to Kant.[19] Hume's own intellectual background will occupy us in Chapter 4. The present point is that even in the demanding sense of 'tradition' that requires not just *similarity* but *influence*, the tradition of pragmatic genealogy can be said to have a notable source in Hume's thought, which is to say right in the ancestral heartland of Anglophone philosophy. Realizing this may help reconcile contemporary analytic philosophy to a method it has long considered alien to itself.

1.2 A Systematic Account of the Method

My second main objective in this book is to develop a systematic account of this pragmatic genealogical method that identifies a rationale for its use in philosophy and makes sense of its many rather baffling features. How does this puzzling hybrid, which is neither a straightforward historical explanation nor a pure

---

may have been influenced in this direction by John Anderson' (Williams 2010, 215n4). Another precursor to Craig's genealogy is Oswald Hanfling's 'structural account' of knowledge (1985).

[18] On Hume's influence on Nietzsche, see Brobjer (2008a, b), Mabille (2009, 73), Emden (2014, 108–9), and Kail (2016). On Hume and Darwin, see Wild (2008, 2011) and Slavov (2019). On Darwin's influence on Nietzsche and its refraction through Rée, see Richardson (2004, 16–17), Small (2005), Janaway (2007, 78), Brobjer (2008a, b), Young (2010, 414), Emden (2014), and Hufendiek (2019). Already in 1872—before he met Rée—Nietzsche described Darwinism as something 'I hold to be true' (*eKGWB*, 1872, 19[132]).

[19] See Drossbach (1884), Fischer (1869, 37–45), Ueberweg (1866, 121–6), and Lange (1866, 145, 237–43, 258–64).

justification, actually work, and why should we bother with its elaborate narratives if we are interested in the point of our current conceptual practices?

These questions only become more acute once one takes a closer look at an example of the method. Take Williams's *Truth and Truthfulness*, whose subtitle introduces it as 'an essay in genealogy'. The genealogy starts out from a 'State of Nature' that is not localized in either space or time. In it, we find 'a small society of human beings, sharing a common language, with no elaborate technology and no form of writing' (2002, 41). These human beings, we are told, need information about their immediate environment if they are to satisfy even their most basic needs. They do not just rely on their five senses to acquire it, but begin to cooperate, in particular by pooling information. They then begin to cultivate the dispositions that make *good* contributors to the pool, and that is where dispositions of accuracy and sincerity, the two components of truthfulness on Williams's view, make their first appearance. Initially, the value of these dispositions is understood in purely instrumental terms, so that they are treated merely as means to the effective sharing of information. But as the genealogy progresses, they come to be regarded as worth exhibiting for their own sake—as *virtues*. These two virtues are still quite unlike what we today mean by accuracy and sincerity, however. For one thing, they only apply to the immediate environment, so that there is no expectation that one should be accurate in speaking about the distant past. They are also not very demanding: sincerity only requires that one come out with one's occurrent beliefs and desires, not with what one *really* believes or desires. To bridge this gap, Williams leaves the fictional state of nature behind and moves into real history—more specifically, into the ancient Greece of Thucydides. There, we are told, accuracy is extended to cover the distant historical past. We then fast-forward to the Romantic period, where sincerity is elaborated into the more demanding virtue of authenticity, though in two different forms: one associated with Rousseau and the other with Diderot. Finally, we move on to consider how truthfulness developed in the context of modern liberal politics, where we are told that truthfulness becomes an important instrument of liberalism 'by serving as the sharp end of a critique of injustice' (2002, 209) and by encouraging truthful history. This genealogical story is not offered merely as an explanation, but is meant to be affirmative or 'vindicatory' (2002, 36): it aims to strengthen our confidence in truthfulness.

This is a caricature of Williams's genealogy, but it gives one a sense of the puzzles with which it has presented even careful readers of *Truth and Truthfulness*: Thomas Nagel (2009b, 134) expresses puzzlement over the project of vindicating through genealogy; Colin Koopman (2013, 20, 64–5, 74, 87) charges Williams with committing the 'genetic fallacy' in conflating genesis and justification; Colin McGinn (2003) finds the genealogical story redundant given that instrumental considerations are supposed to vindicate; Richard Rorty (2002) confesses himself unable to see the connection between the fictional and the historical parts of

the book.[20] And indeed, the way Williams combines insistence on the need for philosophy to involve itself in history with insouciance towards historical detail is puzzling. How can such a breezy romp through history pass as a genealogy of truthfulness? Is Williams perhaps 'trying to press more out of genealogy than is really there' (Koopman 2013, 71) by claiming that his genealogical explanation carries a vindicatory force? And how, in particular, are we to make sense of all this talk of a 'State of Nature'?

Any attempt to assimilate the state-of-nature part of the genealogy to the conjectural but at least professedly historical narratives of evolutionary psychology is soon discouraged. Williams, like Craig before him, is keen to distance his project from evolutionary psychology's conjectural histories and from the criticism they have encountered.[21] The state of nature they start out from, these genealogists insist, is not the Pleistocene. It is not even a conjectural depiction of the distant past or of our environment of evolutionary adaptation. The state of nature is a fiction—there is nothing more to be *found out* about it. In fact, it need not even be so much as possible.[22] As Hume insisted, it is 'a mere philosophical fiction, which never had, and never cou'd have any reality' (*T*, 3.2.2.14).

While this helps demarcate pragmatic genealogy from evolutionary psychology, it seems to do so at the cost of stripping this state-of-nature fiction of all pretensions to doing serious philosophical work. Is not, in Bentham's phrase, the season of fiction now over?[23] How are fictions supposed to tell us anything about reality? In particular, how are situations that are not just counterfactual but even counterpossible supposed to yield insights into the conceptual practices we actually have? And even if they succeed, how are these avowedly fictional stories supposed to mesh with real history? Either one tells a story that tries to be truthful to how things actually developed at some level of description, if only metaphorically, or one abandons this ambition. But Williams's genealogy moves seamlessly from hypothesizing into the blue to deciphering the grey, hieroglyphic writing of the past, as Nietzsche's famous contrast has it (*GM*, P, §7). How can fiction dovetail into history, *mythos* into *logos*, the way Williams moves from the state of nature to Thucydides and thence to Rousseau? Finally, even if these hurdles could be overcome and a coherent and continuous hybrid narrative could be constructed along these lines, setting out from the state of nature seems to undermine the point of telling a genealogy in the first place: it seems to renege on the promise of historicization carried by the term 'genealogy' and to dehistoricize one's object by

---

[20]   More generally, what exactly the book's 'circuitous' (Elgin 2005, 343) argument is supposed to be has been contested. Reactions have ranged from hailing it as 'the most interesting set of reflections on the values of truth and truth-telling in living memory' (Hacking 2004, 137) to questioning whether the book is more than 'a collection of loosely related essays on truth' (Fleischacker 2004, 382). The first monograph on Williams calls it 'a collection of interesting intellectual tributaries feeding a somewhat elusive main channel' (Jenkins 2006, 163).

[21]   See Craig (1990, 10; 2007) and Williams (2002, 27–30). For examples of such criticism of evolutionary psychology, see Dupré (1998, 2001) and Driscoll (2015).

[22]   See Williams (2002, 30).   [23]   See Bentham (1988, 53).

placing it in that 'maximally ahistorical setting' (Fricker 2007, 108–9) that is the state of nature—which, on top of everything, is anyway notorious for providing 'a blank canvas onto which a philosopher may paint the image of his personal theoretical predilections' (Fricker 1998, 164).

All of these worries have considerable force, and it is no surprise that where the approach I have been calling 'pragmatic genealogy' has been noticed at all, it has been considered either entirely *redundant* or severely *restricted* in its scope and power.[24] Those who take pragmatic genealogy to be redundant think that if we want to get at the function of our present practices, a synchronic approach is the shortest route, and genealogical state-of-nature fictions can add nothing but colour to these ascriptions of functionality; and if we are interested not in our present practices, but in some earlier historical form they took, we should do real documentary history instead of contenting ourselves with simplistic and fanciful just-so stories—there is plenty of careful contemporary history that is in the business of uncovering functions.[25] Of course, there are also those who have discerned more merit in the method. But even these more sympathetic interpreters see pragmatic genealogy as restricted in several respects: the state of nature's claim to being explanatory might be salvaged by interpreting it as an abstract depiction of extremely general facts about the human condition, but this restricts the method's scope of application to the explanation of anthropological universals; it also restricts the method's freedom in depicting the state of nature, for when it strays too far from reality, the state of nature loses its explanatory value; and once we have granted pragmatic genealogy entry into the realm of genuine causal explanation by restricting it in this fashion, it runs into the firmly entrenched idea that if genealogical explanations of any kind have a normative upshot at all, they are restricted to subverting claims to being normal, natural, or necessary by revealing the contingency of our arrangements—an idea that would seem to bar pragmatic genealogy from having any effect in the space of reasons, since contingency is precisely what it will *not* reveal if it is restricted to dealing with anthropological universals. We are thus left with the view that pragmatic genealogy must be either redundant or else restricted to unimaginatively saying rather little about the few practices that fall into its proper remit.

Against this, I develop an account of pragmatic genealogy that shows it to be neither redundant nor restricted in any of these ways. Scrutinizing past applications of the method instead reveals pragmatic genealogy to be a powerful and well-motivated elaboration of the synchronic approach. In particular, I argue that when dealing with what I call *self-effacingly functional* practices—practices that are functional *only insofar as* and *because* we do not engage in them for their

---

[24] Proponents of the redundancy view include McGinn (2003), Dutilh Novaes (2015), Hacking (2005, 168), Hannon (2019, 52–3), and Koopman (2009; 2013, 71); proponents of the restriction view include, with qualifications I shall come to in later chapters, Craig (2007, 192–3), Fricker (2007, 114), and Wild (manuscript).

[25] See, e.g., Ogilvie (2019).

functionality—the synchronic approach is prone to miss explanatory connections between the instrumental and the non-instrumental aspects of such practices that pragmatic genealogy is better able to bring out. Second, when dealing with strongly *historically inflected* practices that lack a paradigm case displaying the practice's connection to human needs, the synchronic approach fails to get a grip altogether; by achieving a grip even here, pragmatic genealogy proves a valuable addition to our methodological repertoire. Third, far from being an inferior substitute for real history, pragmatic genealogy presents philosophers with perspicuous representations of practices' relations to needs that are tailored to the demands of relevance, salience, and persuasiveness that are specific to philosophy. Fourth, the method is not restricted to depicting anthropological universals or highly generic facts about the human condition, but can also model local problems deriving from local needs—at its best, pragmatic genealogy offers us a comprehensive view of what a conceptual practice does for us, placing and relating the respects in which the practice answers to both generic and local needs. Fifth, pragmatic genealogy can profitably stray from reality and involve distortions which explain why things are as they are by vividly bringing out how they would fail us if they were different. And sixth, far from being normatively inert, pragmatic genealogy can affect the space of reasons by showing us what reasons we have for or against cultivating a given conceptual practice—although in contrast to the genealogical methods described by Michel Foucault (1971, 1975) and Alasdair MacIntyre (1988), pragmatic genealogy tends to be affirmative or vindicatory before it is destabilizing or subversive: like any broadly functionalist explanation, a pragmatic genealogy begins by showing how something is somehow worthwhile for someone, even if it then continues as a narrative of loss of functionality, or if the *pro tanto* vindication *relative to* some practical concern amounts to an indictment in the eyes of those who have no wish to see that concern satisfied.

To substantiate these claims and to understand what fictional state-of-nature genealogies can be true *to*, we need to question the traditional role allocation according to which the state of nature merely justifies while real history merely explains. We need to make room for the idea that pragmatic genealogies can be genuinely explanatory *despite* operating with a fiction, and that they can be normatively significant *despite* being explanatory.

To this end, we can take our cue from two interpretive strategies that have grown out of attempts to make sense of Craig's genealogy in particular. One is what might be called the *actualist* interpretation advocated by Miranda Fricker (2016b, forthcoming). On this view, temporal priority within a genealogical narrative corresponds in reality to explanatory priority within our *actual* conceptual practices, and the primitive form of a practice considered at the beginning of the genealogy stands for a *paradigm case* of our actual practice which is presented by the genealogy as explanatorily basic. The second strategy is what might be

called the *dynamic model* interpretation advanced by Martin Kusch (2009b, 2011, 2013). On this view, genealogies are not just elaborate ways of describing our actual practices. They involve genuine historicization and fictionalization, because they are best interpreted as *models* which at first involve strong idealization, but are then gradually de-idealized to approximate their target system. They provide models with a time axis—*dynamic* models—explaining why we came to think as we do.

My own approach makes use of both interpretations to develop a general account of this genealogical method that turns on recognizing the extent to which it embodies a certain kind of pragmatism—hence the term 'pragmatic genealogy'. On my account, pragmatic genealogies are an example of philosophy as model-building,[26] and I draw on Kusch's insight that pragmatic genealogies are dynamic models that work through idealization and de-idealization. But Kusch thinks of all the dynamic models that these genealogies present us with as constituting a form of history, albeit history of a particularly abstract kind. For him, they are 'Aristotelian' idealizations that abstract away from the particular without distorting it rather than 'Galilean' idealizations that distort reality in order to illuminate it. On his interpretation, therefore, the order of genealogical development must correspond to the order of historical development. By contrast, I think there is something importantly right about Fricker's insight that the order of priority in these genealogies corresponds in the first instance to explanatory rather than to historical priority. But we need not take this to mean that the genealogies simply describe our present practices. They can be dynamic models involving genuine fictionalization and historicization, only of a kind that is less beholden to history—indeed, I shall argue that part of their power derives from their ability to describe as developing sequentially what in reality had to develop together, so as to help us untangle the array of needs to which some of our most fundamental ideas answer; and part of their power derives from their ability to describe counterfactual or even counterpossible developments in order to help us understand why we think as we do—for instance, to explore genealogical stages that were probably never realized because they are hopelessly unstable (such as truthfulness being valued purely instrumentally), but whose instability tells us something about why we find the slightly more complex stages we really do find in history (such as truthfulness being valued intrinsically).[27]

---

[26] A conception of philosophy as model-building is defended by Paul (2012), for example, who notably advocates the use of fictional situations in metaphysics, and by Williamson (2017; 2018a, ch. 10; 2018b), who advocates the use of formal models in particular, but argues in his 2016 Annual Lecture of the Royal Institute of Philosophy that many areas of philosophy should be thought of as aiming to offer ever better models of highly complex target systems rather than to formulate necessitated universal generalizations and test them against counterexamples.

[27] The view that distorting idealization can enhance rather than impede understanding has been gaining support since the 1980s. See the essays in Grimm, Baumberger, and Ammon (2016) as well as Appiah (2017), Elgin (2007), Strevens (2008, ch. 8), and Weisberg (2007, 2013).

On this interpretation, pragmatic genealogies stand to more regularly historiographical genealogies much as sense-making in terms of practical pressures stands to sense-making in terms of causal-historical processes. Imagine having to explain to someone utterly unfamiliar with our culture why a car has the shape it does. One could do it by enumerating the different stages of the car's actual formation on the assembly line, thereby describing the *causal construction* of the car; or one could explain the design of a finished car as reflecting a series of needs, thereby offering a *pragmatic reconstruction* of the car.[28] Most basically, the design of a car reflects a need for mobility; but it is further determined by the need to see certain practically relevant parts of one's surroundings, the need to stay warm and dry, the need to sit comfortably—and so on, down to the need to follow currently prevailing aesthetic trends. Picture a computer animation starting out from a primitive geometrical shape and gradually reaching something recognizably car-like by successively factoring in the various needs of car-users and warping the shape to meet them. The stages of this formation process would not at all correspond to the steps involved in actually assembling a car. But they would reveal how various aspects of car design reflect and answer to a specific combination of needs. A similar bifurcation is marked by the contrast between primarily historiographical genealogies and pragmatic genealogies: primarily historiographical genealogies can be compared to descriptions of the assembly line, concerned in the first instance to offer an accurate depiction of the stages and forces through which something was actually constructed; the genealogies discussed in this book, by contrast, are better compared to models of needs: they are in the first instance concerned to offer an accurate depiction of the variety of needs that something serves, and hence draw on a succession of practical pressures rather than on a succession of causal-historical forces (the distinction survives the observation that practical pressures are also causal-historical forces, because there is no expectation that the succession of practical pressures must correspond to the succession of causal-historical forces). *Which* practical pressures should figure in the dynamic models is partly a matter of one's interests and purposes in offering the models and partly a matter of one's understanding of psychology, sociology, and history. But while history informs the dynamic models, it is not the primary purpose of these models to mirror actual historical development. Their primary purpose is to extricate from the nit-and-grit of history the main practical pressures and dynamics that have sculpted our conceptual practices and that help us understand their retention, elaboration, and differentiation.

Accordingly, it is an important feature of pragmatic genealogies that they allow genealogists to start out neither from the present nor from a particular point in the past, but from counterfactual situations that bring out the relation of simple

---

[28] See Kappel (2010) for an elaborate example of an explanation along these lines.

prototypes to certain needs. An example of such a genealogy that starts from a probably counterfactual situation is the genealogy of money as told by Philip Pettit on the basis of Carl Menger's (1892) classic work.[29] What does the institution of money do for us, and if it did not exist, why might we invent it? To find out, Pettit invites us to imagine a society in a fictional state of nature—akin to the one he uses for his genealogy of moral desirability and responsibility, which he fittingly calls 'Erewhon', a Butlerian anagram of 'nowhere'—in which money does not exist: a pure barter society, in which commodities are exchanged directly for other commodities. In reality, there has probably never been such a pure barter society (Graeber 2011). But we can usefully start out from there in order to explore the practical pressures that might lead such a society to develop money. In such a society, agents have a pressing second-order need, namely the need to get, by way of exchange, the commodities that they directly need at a given time. If I am an individual in this barter society, however, I face a problem: at any given time, the probability of there being another individual who happens to be able and willing to exchange precisely what I then directly need, and who happens to need precisely what I am able and willing to give in return, is frightfully low, and the probability of our running into each other and performing the exchange in time even lower. The solution lies in the realization that some commodities are easier to exchange than others—they have greater *marketability* (or *Absatzfähigkeit*, as Menger puts it)—because the demand for them is greater, more widespread, and more stable than for other commodities. This means that it is rational for me to exchange my commodities for something I do not directly need if its marketability is greater, for this will increase the probability of my being able to get what I do need by way of exchange. In addition to functioning as commodities in their own right, therefore, highly marketable commodities come to function as *media of exchange*, especially those that also happen to be portable, durable, and divisible, such as precious metals, cowry shells, or, in post-World War II Berlin, cigarettes. And every time a commodity is used as an intermediary in this way, this further increases its marketability by contributing to its wider recognition *as* a medium of exchange. Once one of these commodities reaches the point where it is commonly known to be accepted by everyone, it has become a *general* medium of exchange—in other words, something very like what we call 'money'. The actual historical emergence of money may have owed more to the state's need for a general currency unit for tax payments.[30] But what this genealogy suggests is that even in a pure barter society, something very like money would be nearly bound to arise, because money turns out to discharge a crucial function: it alleviates the fundamental problem of finding suitable exchange partners.

---

[29] See Pettit (2018, 50–2). For a more sustained philosophical discussion of Menger's account, see Tieffenbach (2011).
[30] See Wray (1998, 50) and Karimzadi (2015).

Of course, as this simple prototype of our institution of money illustrates, the dynamic models of cultural development we can construct out of very basic human needs will, for many purposes, still be *too* simple. To bridge the gap between these all-too-simple prototypes and the practices we actually have, the models will need to be further de-idealized in the direction of our actual situation. This can be done by factoring in more *socio-historically local* needs that drive the practice's development further and bring it closer to the practice we recognize as ours. We might tailor the genealogy of money to our own situation, for example, by factoring in the various pressures driving the emergence of certificates or tokens representing commodities and the eventual emancipation of those tokens from the commodities they represent.[31] But although a pragmatic genealogy that gradually factors in the influence of ever more socio-historically local needs is likely to end up modelling history at a highly abstract level, the fact that something happens later in the genealogy does not *necessarily* mean that it happened later in history. *Later* primarily means *less idealized*. The temporal order of the genealogical model is, in the first instance, the order in which the genealogist chooses to take complicating factors into account in order to offer us a maximally surveyable or perspicuous representation—an *übersichtliche Darstellung*, to use the Wittgensteinian term—of the complex tangle of needs at the roots of our conceptual practices.[32]

My second contention, then, is that pragmatic genealogies are best interpreted not as narratives of causal or historical construction, but as narratives of pragmatic reconstruction: they are dynamic models serving to reverse-engineer the functions or points of our conceptual practices in relation to both generic and local needs, thereby enabling us both to explain why we came to think as we do and to evaluate what these conceptual practices do for us. Explanation and evaluation are aims that eventually pull in different directions. But they can be pursued some of the way together, since, as the genealogies we will consider illustrate, many of the practical dynamics at the roots of our ideas are as relevant to their explanation as to their evaluation.

---

[31] See Pettit (2018, 51).
[32] See Owen (2001) for an argument to the effect that in the context of moral and political philosophy, genealogy is one form that Wittgensteinian *übersichtliche Darstellung* may take. In the *Philosophical Investigations*, Wittgenstein writes: 'A surveyable representation produces precisely that kind of understanding which consists in "seeing connections". Hence the importance of finding and inventing *intermediate links*' (2009, §122). For an elucidation of Wittgenstein's notion of *übersichtliche Darstellung*, which he developed out of Goethe's 'morphological method' of using models to organize and understand the development of organisms, see Schulte (2017), who also highlights that the 1931 manuscript version of §122 continued: 'But an hypothetical connecting link should in this case do nothing but direct the attention to the similarity, the relatedness, of the *facts*. As one might illustrate an internal relation of a circle to an ellipse by gradually converting an ellipse into a circle; *but not in order to assert that a certain ellipse actually, historically, had originated from a circle* (evolutionary hypothesis) but only in order to sharpen our eye for a formal connection'.

This interpretation will also bring out the relevance of pragmatic genealogy to philosophers' increasing preoccupation with conceptual engineering.[33] Whereas conceptual engineering involves starting out from what we want a concept to achieve and specifying a concept that will achieve it, conceptual *reverse*-engineering works not *from* the function to the conceptual practice that would perform it, but rather from the conceptual practice *to* the function it performs.[34] Reverse-engineering is thus a backward-looking enterprise that can reveal what our conceptual practices do for us when we do not yet know it. But as we shall see in Chapters 2 and 8, in particular, reverse-engineering nevertheless can and should guide the forward-looking enterprise of engineering better concepts.

1.3 Doing Systematic Philosophy by Doing History of Philosophy

All metaphilosophy runs the risk of having a high-altitude feel to it, of remaining a bloodless abstraction at two removes from reality. Our systematic sketch of the method of pragmatic genealogy and its rationale needs to be fleshed out. One template for this is Descartes's *Discours de la méthode*, where the exposition of a method is followed by treatises on dioptrics and meteors which, as the title page had it, *font essais de cette méthode*—are essays of this method. But in the case of pragmatic genealogy, there already *are* essays of the method—Williams's *Truth and Truthfulness: An Essay in Genealogy* follows the Cartesian model to the letter, opening with methodological reflections in the early chapters and devoting the bulk of the book to showing the method in action. For all that, the method as much as its implementation have remained more than a little obscure. The problem is thus not that essays of the method are lacking, but that the method they instantiate is ill-understood, which in turn leaves it unclear what should or should not count as an instantiation of it. Although I have myself made use of pragmatic genealogy on occasion,[35] I have decided against adding to the list of

---

[33] As evidenced by Cappelen's programmatic monograph on conceptual engineering (2018) and the essays in Burgess, Cappelen, and Plunkett (2020). My interpretation of pragmatic genealogy renders it particularly relevant to approaches seeking to improve conceptual practices on the basis of their function or point, such as Haslanger (2000, 2012a), Brigandt (2010), Brigandt and Rosario (2020), Nado (2019), Simion and Kelp (forthcoming), and Thomasson (2020). Scharp (2013) and Richard (2019, ch. 6, section 6) also conceive of conceptual engineering in terms that invite a focus on functions. See also Plunkett (2016) for a rich discussion of how history, by alerting us to under-appreciated aspects of the content of our concepts and furnishing us with alternative concepts, can inform normative reflection on which concepts to use.

[34] My use of the term 'reverse-engineering' differs from that of Dogramaci (2012) and Hannon (2019, 22–7) in that they use it narrowly to refer to the method of looking at the present usage of words with a view to hypothesizing their function, while I use it more broadly to cover any attempt to explain why we think (and speak) as we do in terms of the practical point of coming to do so.

[35] See Cueni and Queloz (2021).

examples with this book, and resolved instead to provide a unifying interpretation of the examples we have. Fresh perspectives in systematic philosophy also generate fresh perspectives on the history of philosophy: they provide new lenses through which to view the materials of the past, new threads on which to string old ideas, and new frames in which to set them. Pragmatic genealogy is no exception. Scrutinizing the history of philosophy through the methodological lens of pragmatic genealogy reveals under-appreciated aspects and continuities in the works of Hume, Nietzsche, Craig, Williams, Fricker, and, by implication, others whose work will be illuminated by a revised understanding of these authors.[36] Insofar as the genealogies of these authors have been discussed at all, they have typically been discussed separately. By bringing them together in this book, I attempt to situate these authors as proponents of a shared philosophical method with a serious pedigree tracing back to Nietzsche and Hume, and to show how their respective genealogies support each other. The hope is that the end result will be far more satisfying than a defence of any of these authors taken in isolation. By subsuming these various projects under a common terminology and aligning them in a tradition, we get a sense of their commonalities and differences, of the variety of concerns that animate pragmatic genealogy as a method, and of its range, possibilities, and strengths. Fleshing out methodological reflections with antecedent examples of the method in this way keeps methodological theorizing honest: instead of tailoring the illustration to the methodological theory, the methodological theory has to prove itself by fitting its antecedent applications.

This book thus pursues two connected aims: to uncover the methodological tradition of pragmatic genealogy and to make the case for that method by offering an attractive systematic interpretation of it that demonstrates its value to the philosophical enterprise. These two aims are connected, because the book seeks to make the case for the method *by* uncovering a tradition of using it, thereby bridging the systematic–historical divide as well as the analytic–continental divide. Relating systematic philosophical reflection to the history of the subject in this mutually beneficial manner is one way in which the history of philosophy can, in Williams's phrase, be done *philosophically*; that is, in a way that yields philosophy before it yields history.[37] This is particularly true in this case, since on the interpretation I propose, pragmatic genealogies *themselves* yield philosophy before they yield history, and moreover yield the kind of philosophy that helps us make sense of how the ideas of the past relate to those of the present. If Hume or Nietzsche were simply tracing the history of ideas like justice or truthfulness up to their own time, this would leave it unclear how their understanding of these ideas

---

[36] I am thinking in particular of Rousseau (1977), Smith (2002), Hart (1961), and Wittgenstein (2009), though I cannot elaborate on this here; but see Neuhouser (2014), Rasmussen (2017), and Pettit (2019).
[37] See Williams (2006a) and Queloz (2017).

related to our own understanding of them today, and to what extent these were understandings of the same ideas at all. But if, as I shall argue, they are offering dynamic models designed to help us grasp how local elaborations of these ideas are rooted in more widely shared needs, they are providing us with the means to see our ideas and theirs as different elaborations of the same ideas, whose similarities and differences become intelligible as reflecting similarities and differences in our respective needs. In telling their genealogies, the pragmatic genealogists thus themselves build the bridges that connect their ideas to ours.

The next two chapters (Chapters 2 and 3) articulate in more detail what the method of pragmatic genealogy consists in, why one might want to use it, and which types of cases it is particularly well suited for. With the methodological lens of pragmatic genealogy ground and polished by the end of the third chapter—which is both structurally and substantially the pivotal chapter of the book—we can then peer through that lens at the history of philosophy and bring the genealogical explanations of Hume, Nietzsche, Craig, Williams, and Fricker into focus. These historical chapters are not just exegetical, however. Each reconstruction serves to illustrate broader methodological points: Hume demonstrates the functions of the state-of-nature fiction as well as the method's power to bypass circular or overly intellectualist explanations; Nietzsche reveals the value of hardening and sharpening one's functional hypothesizing 'under the hammer-blow of historical knowledge' (*HA*, I, §37), as he characteristically put it, while also alerting us to the need to be more sensitive than perhaps Hume was to the possibility that something originally useful might deteriorate into dysfunctionality; Craig shows us the pressures driving concepts to shed the traces of their origins in subjective needs; Williams brings out the pressures driving their further de-instrumentalization into concepts of intrinsic values; and Fricker indicates how the state-of-nature fiction can be politicized and used for ameliorative purposes. We then return to a more purely systematic perspective and address a battery of objections in the last two chapters (Chapters 9 and 10), clarifying what the normative significance of pragmatic genealogies consists in, how the notion of having a point should be understood, what the role of needs is, and how genealogy offers us a concrete model for the pursuit of philosophy as a humanistic discipline. The hope is that this combination of systematic and historical perspectives yields a nuanced understanding not only of these pragmatic genealogists, but also of the method and its still largely untapped potential.

2

The Benefits of Reverse-Engineering

Why should we bother to reverse-engineer the points of ideas? My aim in this chapter is to highlight some of the benefits that genealogy as conceptual reverse-engineering promises to deliver. After working through seven things that reverse-engineering can do for us, I focus on three benefits that reverse-engineering through genealogy is particularly well suited to delivering: far from issuing in the kind of reductively instrumental view of things often associated with naturalism, pragmatism, and genealogy, the method offers us *explanation without reduction*, combining naturalism and pragmatism into a non-reductive framework that can help us understand what led our ideas to shed the traces of their practical origins; far from being normatively inert in the way that genealogical explanations are commonly taken to be, pragmatic genealogy can *affect the space of reasons*, subverting or vindicating our ideas by weakening or strengthening our confidence in them; and far from being boringly conservative when its upshot is vindicatory, pragmatic genealogy can facilitate *responsible conceptual engineering* by alerting us to the multiplicity of functions we need to take into account as we revise our conceptual practices. Finally, I situate pragmatic genealogy in a broader methodological landscape before turning, in Chapter 3, to the question of when it is to be preferred over non-genealogical forms of reverse-engineering.

2.1 From a Practical Point of View

Looking at ideas from a practical point of view can induce a striking gestalt switch in the way we see them, because we often do not think of them in terms of their point. Part of the reason is the phenomenon familiar to designers and architects that while bad design has a way of calling attention to itself, the mark of good design is often its invisibility, leaving it at risk of remaining unappreciated. More generally, a prime manifestation of the fact that something *has* a point is that the question of its point does not arise. If anything leads the question 'What's the point?' to break into consciousness after all, it is typically the dim sense that it has ceased to have an answer.

But with ideas in particular, there is a tendency to think of them as *nothing but* ideas: as ethereal, causally inert things that have no effects, much less a point. Hence, those who trust in the power of ideas frequently encounter the suspicion that mere ideas have no real-world effects. These contrasting attitudes are

personified, for example, in the 1932 encounter between the young Isaiah Berlin, a philosopher and historian of ideas who strongly believed in their power to make a difference,[1] and the already well-established historian Lewis Namier, whose name became a byword for scepticism about the relevance of ideas, and who told Berlin that historians of ideas were 'the least useful kind of historian', because ideas were 'mere interpretations by the mind of deep-seated drives and motives' (Berlin 2014, 127). A scintillating riposte on behalf of ideas had been given a century earlier, when the Victorian writer and historian of the French Revolution Thomas Carlyle faced similar scepticism from a hard-headed businessman who exclaimed: 'Ideas, Mr. Carlyle, ideas, nothing but ideas!' Carlyle retorted: 'There was once a man called Rousseau who wrote a book containing nothing but ideas. The second edition was bound in the skins of those who laughed at the first'.[2]

As Carlyle reminds us, ideas have very real effects—indeed, often unsurveyably many. This is why talk of 'the point' or 'the function' of an idea can be helpful as a way of highlighting those among its effects that we care about for given purposes. The focus of the pragmatic genealogists I discuss lies on effects that are conducive to the satisfaction of human needs. In ascribing a *point* to an idea, they explicitly represent it as tying in with some need in virtue of its tendency to produce effects satisfying that need. This allows them to single out salient because much-needed practical effects of ideas on the lives of those who live by them (I elaborate on this 'need-satisfaction account' of pointfulness in Chapter 9).[3] Talk of '*the* point' of an idea is not meant to foreclose the possibility that it serves multiple points, but to draw attention to what, for given purposes, emerges as the *most salient useful difference* the idea makes, where usefulness, in the genealogies that will concern us, is cashed out primarily in terms of the tendency to satisfy genuine needs. To inquire into the point of an idea is to seek to understand what it *adds*, and consequently also what would be *lost* were the idea to be abandoned.

The telling of such a reverse-engineering story itself only has a point when the audience of the story lacks a clear grasp of whether an idea has a point at all, what needs give point to it, or what further and less obvious points it might serve. Some ideas, such as the revolutionary political ideas of Rousseau's *Social Contract* that Carlyle was alluding to, carry their ambition to have real-world effects on their sleeve to a greater degree than others, and while it can be worth asking whether an

---

[1] Margalit (2013) highlights numerous facets of Berlin's work that illustrate his belief in the power of ideas. That belief also found expression in 'Two Concepts of Liberty': 'Over a hundred years ago, the German poet Heine warned the French not to underestimate the power of ideas: philosophical concepts nurtured in the stillness of a professor's study could destroy a civilisation' (2002, 167).

[2] MacIntyre (1998, 117).

[3] The phrase 'the point of a concept' is used in a variety of ways. My concern here is with the *practical* point: the salient useful difference that living by a concept makes to the lives of concept-users. In Queloz (2019), I distinguish the practical from three other types of points: the *evaluative* point, which is what one has to grasp in order to master a thick concept; the *animating* point, which is the aim or goal one has in mind in applying a concept; and the *inferential* point, which is a salient inferential consequence of the applicability of a concept.

idea really serves the point it *seems* to serve,[4] there are also ideas for which reverse-engineering must prove uninformative. Just as some ideas—*broom* or *screwdriver*—are transparently ideas *of* tools, some ideas are transparently *themselves* tools. Take the idea of a *meeting point*. Nothing beyond that idea needs to come into existence for it to serve its point. It is the idea itself that helps us coordinate by allowing us to think of some place *as* a meeting point. In this case, that observation is not terribly informative, since we likely already thought of that idea as a coordination tool.

For many ideas, however, awareness of their point clearly does not form part of what is involved in living by them. Some ideas even actively direct attention away from practical considerations: the realization that lofty ideas like *truth*, *knowledge*, or *justice* might have anything to do with something as mundane as human needs is one that these ideas themselves conceal by inviting us to think in altogether less instrumental terms. As a result, these are ideas we might need without knowing that we need them. Needs are not necessarily felt needs. Particularly where ideas discourage or even resist being understood in instrumental terms, therefore, reverse-engineering an idea's practical origins in certain needs can be informative, providing a fresh perspective on it and putting one in a position to ask further questions, such as: are these needs we now share? Whose needs exactly are they? Are they needs we endorse?

But what exactly is gained by having a sense of an idea's point and being in a position to ask these questions? There are a number of benefits, some of which are not exclusive to pragmatic genealogy, but are shared by non-genealogical forms of reverse-engineering. When reverse-engineering should be pursued genealogically will be the focus of Chapter 3. For now, let us look at why one might want to pursue it at all by considering seven general virtues of conceptual reverse-engineering.

2.2 Seven Virtues of Conceptual Reverse-Engineering

In spelling out what conceptual reverse-engineering can do for us, it will be useful to draw on some of the technical vocabulary that has sprung up around the notion of a 'concept'. But since different disciplines and subdisciplines have—no doubt for good practical reasons—come to understand rather different things by the term 'concept',[5] it may help to start with a working characterization of what I shall

---

[4] This would turn pragmatic genealogy into an instrument of ideology critique. See Anghie (2007, 292), Moyn (2010), Srinivasan (2019, 142), and Queloz (2019, §5) for critiques along those lines. For accounts of the critical force of history in general, see Epstein (2010) and Cueni and Queloz (manuscript).

[5] Concepts have variously been taken to be *mental representations* (Fodor 1998, 2004), *abstract entities* such as Fregean senses (Chalmers 2011; Peacocke 1992), or *abilities* (Brandom 1994; Dummett

mean by it. For the purposes of this book, I view concepts primarily as thinking techniques: as the norm-governed patterns according to which we move from perception to thought, from thought to thought, and from thought to action. In acquiring concepts, we learn to interpret and evaluate the world around us by forming bundles of dispositions to partition the world in certain ways and draw certain inferences about it, both in our minds and through our actions. To say that I have acquired the concept *tiger*, for example, is to say that I learned reliably to partition the world into tigers and non-tigers, and to infer from something's being a tiger that it is a carnivore and that if I bump into a free-roaming one, I had better run. What concepts I have is not just a matter of what I can articulate, but also of how my capacities for attention, categorization, interpretation, memory, and affect are marshalled and organized in response to different kinds of situations (Haslanger forthcoming, §2). Concepts can be more or less fine-grained and more or less contentful depending on how many partitions and inferences they involve. The bundles of dispositions involved in having a concept possess a normative dimension—they can be manifested correctly or incorrectly. This distinguishes them from *mere* differential dispositions, such as the disposition to laugh whenever tickled. A concept is thus a normative nexus between certain input and output conditions: its input conditions are the *criteria* guiding concept application—the circumstances whose presence licenses the application of the concept—and its output conditions are the *consequences* of concept application— the inferences that can properly be drawn when the concept applies. The norms of differential and inferential behaviour that concept possession involves being sensitive to are the grooves that guide our thoughts and actions by determining what counts as a reason for what—grooves along which our thoughts and actions *actually* run whenever we apply our concepts correctly.[6] Many concepts are dignified with dedicated words by which to express and refer to them, such as the word 'tiger'. But, importantly for the genealogical perspective on concepts, one can also have a concept *before* one has a dedicated word by which to express it.[7] Concepts and words will often enter the genealogical story at different stages: conceptual behaviour (i.e. behaviour structured by concepts) may precede the linguistic means of expressing it, which means may themselves derive their point from the antecedent ubiquity of the concepts they help express; conversely, the

---

1996; Millikan 2000, 2017), for example. For thorough overviews of the different accounts of concepts on offer, see Margolis and Laurence (1999, 2015, 2019). See also Glock (2006a, 2009, 2010) and Haslanger (2012a, 2019, 2020). I shall follow the abilities view in its emphasis on dispositions; unlike those who advocate that view over its main rivals, however, I am not here concerned to deny that these abilities are best explained by appeal to mental representations or abstract entities.

[6] Contrast this normative notion of concepts with the descriptive one, characteristic of work in psychology, on which concepts are 'the tracks our minds prefer to travel on' (Machery 2017, 222).

[7] Sawyer (2020) explores further rationales for distinguishing between thought and talk.

point of introducing a word may be to catalyse the emergence of conceptual behaviour that is as yet largely absent and that one would like to promote.

In asking after the points of concepts, then, we are asking after the points of particular bundles of differential and inferential dispositions. This puts the focus on what might be called the *business end* of concepts—not on the way in which concepts are encoded in the brain or represented in the mind, or on the norms or proprieties of use that a competent concept-user is sensitive to, but on what the *possession* of just these bundles of dispositions, sensitive to just these proprieties of use, allows us to *achieve*.[8] This focus on the practical side of things is reflected in the expression 'conceptual practice', where (as we saw in Chapter 1) a conceptual practice is defined as a community's practice of living by a given concept—of being engaged users of that concept who, in the conduct of their own affairs, prove sensitive to the distinctions encoded in the concept and responsive to the reasons that the concept's applicability generates. If we see a conceptual practice as a technique that renders concept-users sensitive to certain features of the world and links them in their minds with certain inferences in thought and action, we can ask: what is the point of living by a concept that tracks just these features and links them with just those inferences?

I take it that this working characterization of concepts and conceptual practices can ultimately be rendered compatible with most leading accounts of the nature of concepts, and that nothing in what follows depends on a prior commitment to a more specific account of concepts.[9] This is not to say that which ontological categories one uses in explicating the nature of concepts is irrelevant to their function. Abilities, mental representations, and abstract objects really are different things that will be recruited to human ends in subtly different ways for subtly different purposes. The reason no particular view is presupposed in this book is that it does not start from some preconceived view of what concepts are that is then imposed top-down on various subject matters, but rather lets the appropriate set of theoretical notions grow out of each inquiry into practical origins. It is left to the practical problem at hand to determine what kind of response is certainly called for—what one must minimally be able to *do* or *have* in order to solve the problem. Of course, understanding the practical difference made by a single concept often leads one to consider an entire *complex* of interrelated dispositions, concepts, evaluative attitudes, and practices, with each component contributing to

---

[8] The same abilities may be realized in different individuals in different ways (Haslanger forthcoming, §2). How one then spells out the conditions of concept possession is a matter of some debate: see Boghossian (2003), Williamson (2003), Fodor (2004), Glock (2006a, 2009, 2010), Eklund (2007), and Scharp (2013, ch. 2).

[9] An account that seems at first particularly well suited to a genealogical approach is Sainsbury and Tye's (2012) originalist theory of concepts; but their focus lies on showing how various puzzles in the philosophy of language can be solved if we treat the reference of concepts as fixed by their original historical use and individuate concepts by those origins. Pragmatic genealogy, by contrast, focuses on practical rather than historical origins and invites us to individuate concepts by their point.

the overall efficacy of the bundle. But this does not bar us from telling apart the contributions of the various components. We can still distinguish between what the emergence of a discriminatory disposition adds and what the concept *of* that disposition adds, for example, or identify what is gained by turning an originally valuation-free concept into an evaluative concept, even if these elements have to arise together to make a difference.

With this working characterization of concepts in place, we can draw out the benefits of conceptual reverse-engineering by comparing it to the more traditional method of conceptual analysis. Conceptual analysis is commonly understood to aim to provide an explicit intension to be measured against the intuitive extension of a given concept.[10] The *extension* of a concept is the set of all cases to which the concept applies. We can determine what a concept's extension is by presenting competent speakers with actual or imagined situations and eliciting, for each situation, their intuitive response as to whether the concept applies in it or not. Using this method, one might in principle sort all imaginable cases into two kinds: those in which the concept applies, and those in which it does not. The first set of cases constitutes the concept's intuitive extension. This intuitive extension can then serve as a measure of success for the articulation of the concept's *intension*: if we have mastered a concept to the point where we can say, for most situations and with some assurance, whether the concept applies in it or not, then that in virtue of which we are able to do this is the concept's intension. The conceptual analyst's task is to render this implicit intension explicit, in the form, for instance, of a list of characteristics whose presence or absence guides our application of the concept. The most stringent form of conceptual analysis aims at a *strict* definition of the concept of X—a definition in terms of individually necessary and jointly sufficient conditions for something to be X.

Where this kind of stringent conceptual analysis runs into difficulty, however, is with conceptual practices that are *internally diverse*. When conceptual practices are held together by criss-crossing relations of family-resemblance rather than a common core (as Wittgenstein suggested was the case with our concept *game*), boiling them down to their highest common factor through conceptual analyses aiming at strict definitions is the wrong approach to take. It is likely to leave us either with a definition that is too thin to be informative, or with no definition at all. Any feature that is not strictly a necessary condition will eventually fall prey to counterexamples and drop out of the final analysis. Craig offers the example of the epistemologist who is at first impressed by the observation that those who *know* that *p* also *believe* that *p*, but then encounters a single counterexample and promptly strikes the belief-condition from the final analysis of knowledge. The observation that *nearly all* knowers are believers is cast aside as irrelevant. 'Of all

---

[10] See Craig (1990, 5, 15–16; 1993).

its deep centrality', Craig laments, 'nothing whatever remains—it could be as incidental as the fact that nearly all knowers are less than 150 years old' (1990, 14). Yet surely the persistent association of knowledge with belief, if real,[11] should tell us something? And how can we be so sure that all and only instances of the concept in fact share a common set of features? The ambition of conceptual analysis has traditionally been to arrive at some necessitated universal generalization of the form: 'necessarily, all *F*s are *G*', such as 'knowledge is justified true belief'. But if the conceptual practice at issue is internally diverse, then that ambition is misplaced. Perhaps the concept just tracks a variety of conditions that are nearly always satisfied while flexibly accommodating the fact that any one condition may be missing under exceptional circumstances.

The tendency to treat counterexamples as decisive would then be equally misplaced. It is only as long as we take philosophy to aim at law-like generalizations that it makes sense to focus on coming up with real or imagined counterexamples: a possible scenario in which a necessitated universal generalization fails to hold can indeed suffice to falsify it. But as I suggested in Chapter 1, pragmatic genealogists are better thought of as model-builders. And if Timothy Williamson (2017, 2018a, b) is right and a lot of philosophy in fact aims to provide models rather than law-like generalizations, then we are underselling philosophy if we measure its success by the number of laws it succeeds in formulating, and we ought not to grant isolated counterexamples the decisive power they have been granted in the past.[12] As an economist once pointed out, Edmund Gettier's (1963) notorious counterexamples to the analysis of knowledge as justified true belief would never have been published in economics, because in a discipline that is self-consciously concerned with idealization and model-building, nobody expects a model to fit each and every case, and the existence of counterexamples would not be considered newsworthy (Williamson 2017, 168). The way to supersede a model is not to find a case where it does not hold, but to build a *better* model: a model that improves on the previous one along some dimension—by retaining past explanatory successes while adding new ones, for example.

Even where law-like generalizations and definitions in terms of necessary and sufficient conditions *are* available, moreover, the features that are illuminating for the purposes of philosophy may not all be among the necessary conditions.[13] Why

---

[11] Some have maintained that knowing that *p* in fact *excludes* believing that *p*. See, e.g., Prichard (1950, 86–8).

[12] As Williamson argues, a method that encourages us to abandon a hypothesis already in the face of a single counterexample will be *error-fragile*: a single mistake—such as misidentifying something as a counterexample when it is in fact consistent with the hypothesis—can lead us astray. Philosophy as model-building 'is much less error-fragile, for it gives no such decisive power to a single judgment' (Williamson 2018b, 22).

[13] See Craig (1990, 5, 15–16) and Fricker (2016b, 166). Huddleston (2018) finds the point already in Nietzsche: even if a given concept were to prove definable in terms of a shared denominator, the shared denominator may not be very informative, and it is the variable aspects of concepts that reward attention.

something persists, how it functions, or what its value is are questions that may best be answered by drawing on features that—though characteristic, distinctive, or typical—are not invariably present. We will then profitably widen our gaze to consider the variety of conditions under which something *typically* serves its point. For such conditions—*typical conditions* for short—to possess explanatory value, the concept need not be functional whenever these conditions are given; nor need all or indeed any of the conditions be given whenever the concept is functional.

This is where conceptual reverse-engineering offers a welcome alternative. Its two outstanding virtues as a method are, first, that it is more tolerant of exceptions and counterexamples than explanations in terms of necessary conditions, so that it gets a grip on the diversity that frustrates attempts at conceptual analysis; and second, that it makes explanatory use of merely typical conditions, which in turn allows us to make sense of the diversity frustrating conceptual analysis by showing that this diversity itself serves a function: it reflects a corresponding diversity in the properties that are *typically worth tracking* if such-and-such needs are to be satisfied in the kind of world in which we operate.

Where successful, conceptual reverse-engineering can bring several further benefits. A third one is that organizing an internally diverse conceptual practice in the light of its overarching point *imposes order* on a seemingly messy or disunified practice: it presents some form of the practice as explanatorily basic and renders other forms intelligible as elaborations of that basic form. We may wonder, for instance, why such diametrically opposed phenomena as *earned* forgiveness and *gifted* forgiveness should both be forms of one and the same conceptual practice.[14] Reverse-engineering the point of forgiveness yields an ordering principle that reveals the functional unity underlying the heteroclitic practice. This turns apparent chaos into what Miranda Fricker calls an 'ordered pluralism' (forthcoming). This basic form acts as a beacon around which other forms can be arranged and in relation to which their own derivative point becomes apparent. Variety that was initially baffling is shown to be subservient to a single overarching point. The same overarching point also helps us individuate the concepts whose genealogies we want to trace. Instead of following words through history and relying on homonymy to individuate our object of investigation, we trace lines of genealogical descent in light of functions and functional differentiations. In Hume's phrase, we go by whether conceptual practices 'point all to a like end' (*EPM*, 3.2).

Fourth, by showing how a concept's internal diversity reflects the variety of properties worth tracking in order to serve concept-users' needs, conceptual reverse-engineering can also *resolve standoffs* in philosophical debates over how

---

[14] See Fricker (forthcoming).

a concept should be understood. Where competing analyses of a concept are on offer, reverse-engineering the point of the concept can help us *assess* the different analyses according to whether they fit our functional hypothesis; but it also gives us the means to attempt a *synthesis* of competing analyses—resulting, at the limit, in the complete dissolution of the debate. In the light of a functional hypothesis about the concept's point, it may turn out that it makes most sense for the concept to track *all* the features variously singled out by the different analyses—in which case the mistake was to assume that we had to decide between those analyses.

Fifth, an explanation of why certain needs would naturally lead us to go in for certain conceptual practices can also perform a *demystifying* role: it can dispel the air of mystery about the entities referred to in those practices. By translating venerable but suspiciously empyrean ideas of justice or truth back into nature, reverse-engineering can rid us of the need for metaphysical explanations of the sort that philosophers from Hume through Nietzsche to Rorty and Huw Price take issue with as involving unnecessary ontological commitments or explanatory material. This is most explicit in Pettit's *The Birth of Ethics*. 'The idea', Pettit writes, 'is to demystify our ethical concepts naturalistically by explaining how corresponding concepts might have emerged in a naturalistically unproblematic way' (2018, 54). This demystifying effect of reverse-engineering can be targeted at two kinds of account. On the one hand, it can be targeted at *metaphysical* accounts which seek to make sense of some philosophically puzzling items by introducing special explanatory material. An example is F. P. Ramsey's reverse-engineering of probability statements in response to J. M. Keynes.[15] Puzzled about the nature of probability, Keynes suggested that probability was concerned with objective and unanalysable relations between propositions. Ramsey disarmingly objected that he did not perceive such relations, and suspected others did not perceive them either (1990, 57). Ramsey proposed to look instead at the point of probability statements, which he thought was to express one's confidence or degree of belief in the occurrence of events in order to align one's degrees of belief with those of others and with the frequencies of events so as to facilitate successful action. Probability, on this view, 'is a measurement of belief *qua* basis of action' (1990, 67). Ramsey demystified probability by relating it to human needs and actions rather than to mysterious objective relations.

On the other hand, the demystifying effect of reverse-engineering can be targeted at *eliminativist* accounts which seek to do away with items *because* they seem to demand special explanatory material. Demystification is then a form of rehabilitation in the face of eliminativist reactions to philosophical vexation. An example is Williams's response to Rorty (the topic of Chapter 7): Rorty suggests we do away with the naturalistically suspect intrinsic value of truth, and in

---

[15] See Misak (2016, 175–8) for a fuller discussion.

response, Williams offers a practical derivation of the need to value the truth intrinsically from uncontroversial human needs. Valuing the truth intrinsically is shown to be naturalistically unexceptionable—neither fetish nor superstition, but a way of thinking that naturalists should be comfortable with, as it is firmly rooted in the soil of human concerns.

A sixth benefit or virtue of conceptual reverse-engineering is that it yields a holistic understanding of concepts as pragmatically situated by relating them to contingent facts about concept-users and their circumstances. It thereby exploits the Wittgensteinian insight that careful scrutiny of a concept alone is not going to tell us everything worth knowing about it, and that we must look beyond the concept to the contingent facts that explain its formation and give it its point. As Wittgenstein puts it, 'a natural foundation for the way [a] concept is formed is the complex nature and the variety of human contingencies' (1981, §439), and this suggests that we should be interested in

> the correspondence between concepts and very general facts of nature. (Such facts as mostly do not strike us because of their generality.) But our interest is not thereby thrown back on to these possible causes of concept formation; we are not doing natural science; nor yet natural history—since we can also invent fictitious natural history for our purposes.   (2009, II, §365)

A remarkably similar line is taken by P. F. Strawson when he describes the strand of philosophy that attempts to

> explain, not just how our concepts and types of discourse operate, but why it is that we have such concepts and types of discourse as we do; and what alternatives there might be. This is not an historical enquiry. It attempts to show the natural foundations of our logical, conceptual apparatus, in the way things happen in the world, and in our own natures. . . . It might reasonably be maintained, or ruled, that full understanding of a concept is not achieved until this kind of enquiry is added to the activities of comparing, contrasting and distinguishing.
> (1963, 515–16)[16]

Pursuing this line, conceptual reverse-engineering takes not only contingent facts about the extension into account, but also facts about the needs and capacities of concept-users and their environment in order to achieve a more *holistic* understanding of why a concept with such an intension and extension is worth having. Where conceptual analysis *zooms in* on a concept and tries to analyse its intension

---

[16] Strawson explicitly endorses the use of models to this end, for 'to understand the foundation of our concepts in natural facts, and to envisage alternative possibilities, it is not enough to have a sharp eye for linguistic actualities' (1963, 517).

into its basic constituents, reverse-engineering *zooms out* to bring into view the broader patterns and purposes of human behaviour and the weave of life in which they are embedded. We need a holistic understanding of what the context in which we put our concepts to use is actually like, and how this renders it more useful to operate with certain concepts rather than others, in order to see what our concepts do for us. By thus putting the dependence of concepts on contingent facts centre-stage, reverse-engineering displays what Huw Price calls 'sensitivity to the contingent dependencies of language' (2011, 12).

Finally, conceptual reverse-engineering, though itself a backward-looking enterprise, can serve as a guide and justification for the forward-looking enterprise of conceptual engineering. Even vindicatory genealogy is indicatory of what we can do to improve our concepts: most basically, it puts us in a position to ask how a concept might serve the same point better; but it can also help us see how a concept might need to be adapted in order to recreate the same functionality in a different context, thereby guiding the kind of conceptual innovation that hinges on maintaining continuity of function across very different contexts of application;[17] and it can even show us that a conceptual practice we lack would respond to a very basic and near-universal need we in fact have (see the discussion of Fricker's genealogy in Chapter 8). In all these ways, conceptual reverse-engineering can indicate in what respects and in which directions we have reason to alter our practices, thereby also providing a justification for altering them.

In sum, we have so far identified seven virtues of conceptual reverse-engineering: (i) it is more tolerant of exceptions and counterexamples than explanations in terms of necessary conditions, so that it gets a grip on the diversity that frustrates attempts at conceptual analysis; (ii) it makes explanatory use of typical conditions, which enables it to make sense of the diversity frustrating conceptual analysis by showing that this diversity serves a function, reflecting a corresponding diversity in the properties that are typically worth tracking; (iii) it imposes order on internally diverse practices by identifying certain forms as explanatorily basic and other forms as derivative or elaborative of these basic forms; (iv) it dissolves debates over which account of a concept is the correct one by presenting competing accounts as complementary descriptions of the variety of conditions in which a concept typically fulfils its function; (v) it plays a demystifying role by making us comfortable with puzzling notions that invite metaphysical or eliminativist treatment; (vi) it yields a holistic understanding of concepts as pragmatically situated by relating them to contingent facts

---

[17] See Cueni (2020) for an application of this idea to public law analogies in international legal theory, and Queloz (forthcoming-a) for a discussion of Nietzsche's aspiration to recreate the functionality of Renaissance conceptions of virtue in the modern world.

about concept-users and their circumstances; and (vii) it can guide and justify conceptual engineering.[18]

All of these benefits arguably accrue to conceptual reverse-engineering in general, whether pursued genealogically or not. But there are three further benefits that pragmatic genealogy is particularly well suited to delivering. On the account I propose, pragmatic genealogy offers us explanation without reduction, it affects the space of reasons by weakening or strengthening our confidence in the ideas we live by, and it facilitates responsible conceptual engineering. Let us take them in turn.

2.3 Explanation Without Reduction

Pragmatic genealogy combines three approaches—genealogy, naturalism, and pragmatism—which, deservedly or not, share a reputation for being reductive in the sense of explaining *away* their object, unmasking the higher as really being an instance of the lower: selfishness masquerading as selflessness, instrumental value masquerading as intrinsic value, causes masquerading as reasons, or contingency masquerading as necessity.[19] There is no doubt that these approaches have sometimes been used to this effect, and it is these deflationary uses that Robert Brandom has in mind when he characterizes genealogy as 'the revenge of Enlightenment *naturalism* on Enlightenment *rationalism*', which dispels 'the pretensions of reason' (2015, 3).[20] Suitably understood, however, naturalism and pragmatism can be welded into a non-reductive genealogical method which does not seek to identify one thing with another, but rather helps us understand why a genuinely new thing would arise—it explains without explaining away.

2.3.1 Naturalism

The pragmatic genealogists discussed in this book are all committed to some form of *naturalism*, and this is no coincidence. Hume makes it a methodological precept of his to explain human behaviour wherever possible in terms of mechanisms that are applicable to the rest of the animal kingdom as well: 'When any

---

[18] Hannon (2019, 14–15) highlights other benefits of the approach which I do not pursue here, such as the fact that it helps us adjudicate conflicts of intuitions, or avoid 'verbal disputes' in Chalmers's (2011) sense by clearly identifying the function that a term is to perform.

[19] See Saar (2007), Brandom (2015), and Brassier (2016) for an overview.

[20] See Brandom (2019b, ch. 15, §4) for a critical discussion of reductively naturalistic genealogical explanations as one-sidedly and, if applied globally, self-refutingly *niederträchtig*, debunking normative attitudes by exhaustively explaining them in terms of causes that do not provide reasons for them. In this Hegelian terminology, my concern is rather with genealogical explanations that are *edelmütig*, rationalizing their object instead of debunking it.

hypothesis ... is advanc'd to explain a mental operation, which is common to men and beasts, we must apply the same hypothesis to both' (*T*, 1.3.16.3). Accordingly, the concepts of sense impressions, ideas, associations, and habits that form the nuts and bolts of his project in the *Treatise* are drawn from the animal physiology of his time.[21] The same ambition 'to *naturalize* humanity' (*GS*, §109) is voiced by Nietzsche, albeit in a style that is inimitably his own, when he declares that he aims to

> translate humanity back into nature; ... to make sure that, from now on, humans will stand before humans just as they already stand before the *rest* of nature today, hardened by the discipline of science, with unflinching Oedipus eyes and sealed Odysseus ears, deaf to the lures of the old metaphysical bird catchers who have been whistling to them for far too long: 'You are more! You are higher! You are of different origin!'   (*BGE*, §230)[22]

Similarly, Craig makes a point of noting that his project 'can claim membership' of the 'tradition of naturalism, in which thinkers see man, his behaviour and institutions, as natural facts to be understood as the (broadly speaking causal) outcome of other natural facts' (1990, 8–9). Williams describes his genealogical inquiry as committed to a 'naturalistic outlook' (2002, 60) within which the guiding question is: 'how does the phenomenon in question intelligibly relate to the *rest* of nature, and how, in particular, might it have come about?' (2002, 60). If we can sketch some story—even an impossible story—that answers this question while managing to do without certain resources, this will 'suggest that there could be some actual and much less tidy account which also did without them, and called on no more input' (2000, 157). This is a form of naturalism which even Fricker, though contributing to debates in which 'to naturalize' tends to mean 'to treat as naturally given what is in fact socially constructed',[23] gives every indication of endorsing (Fricker 2007, 112, 118, 129).

   These five philosophers pursue a genealogical naturalism insofar as they try to determine whether we can explain ideas that seem to call for extra explanatory material (such as Divine Commands, Platonic Forms, special faculties of moral intuition, or innate sensitivities) in terms that are as far as possible antecedent to these ideas and the motives bound up with them, but not necessarily antecedent to other human motives of the kind that an experienced and unoptimistic interpreter would discern. This naturalistic stance, which I shall characterize in

---

[21] See Kail (2007a, b, 203), Buckner (2013), and Wild (2014).

[22] Translations are my own, though I consulted the English-language editions listed in the Bibliography. Kail (2016) labels both Hume and Nietzsche 'genealogical naturalists'. See also Clark (1998, xxii), Leiter (2015), and Owen (2007, 6) for discussions of how Nietzsche's genealogical method serves his naturalistic aims.

[23] See Fricker (2007, 137).

more detail in Chapter 5, does not seek to identify the higher with the lower; but it does involve what Williams calls 'an appropriately suspicious rule of method' (1995d, 204), namely that we should never explain a phenomenon in terms of something special to it if we can explain it in terms that we have reason to use anyway elsewhere.

2.3.2 Pragmatism

Deeper even than the pragmatic genealogists' commitment to naturalism is their commitment to a certain form of *pragmatism*.[24] They are writing downstream of the advent of probabilistic explanation (of which Hume is a well-known pioneer within philosophy), and their pragmatism pursues the explanatory strategy of replacing strict *nomological necessity*, which obtains only if something *has* to happen because it is (conditionally) required by inviolable laws, with a more tolerant form of *practical necessity*, which obtains already if something is *highly likely* to happen because it is rendered nearly indispensable by practical pressures.[25] The pragmatic genealogists aim to identify a sense in which, for all their contingency, some concepts have grown out of necessity. In tracing the seemingly transcendent to its roots in human needs, pragmatic genealogy substitutes philosophical anthropology for metaphysics in a way that is characteristic of pragmatism: it reveals how, as William James put it, the trail of the human serpent lies over everything (1978, 37). The pragmatism at work in pragmatic genealogy is not pragmatism as a theory of meaning, however—precisely not: it is an *explanatory* pragmatism accounting for the fact that pragmatism largely fails as a theory of meaning.

The form of pragmatism at work in pragmatic genealogy is best captured by the label *Cambridge pragmatism*, which Cheryl Misak's (2016) book of the same title associates in particular with Peirce, Ramsey, and Wittgenstein.[26] Huw Price and

---

[24] 'Pragmatism', a term first coined by Peirce, has gone through so many hands that it is in danger of losing its embossing. This was the case already in Peirce's lifetime. Seeing the term appropriated by William James, Peirce re-christened his own position 'pragmaticism' in 1905, a name he deemed 'ugly enough to be safe from kidnappers' (1931, 5.414). See Menand (2001) and Misak (2013, 2016) for historical overviews.

[25] See Brandom (2002, 2004, 2019a) for discussions of how this strategy comes to fruition in the 'statistical' nineteenth century, when accounts in terms of natural selection and statistical likelihood were offered to show 'how observed order can arise, contingently, but explicably, out of chaos—as the cumulative diachronic and synchronic result respectively of individually random occurrences' (Brandom 2004, 2); see also Strevens (2013).

[26] See Misak (2016) and Price (2017). See also Blackburn (2013b, 71; 2017), Glock (2017b), Lillehammer (2017), and M. Williams (2013, 128) for discussions of this form of pragmatism. Haslanger's (2012b) distinctive way of fusing pragmatist and genealogical insights is also germane to pragmatic genealogy. I discuss her proposal in Chapter 8.

Simon Blackburn offer characterizations of Cambridge pragmatism that point to its connection with genealogy:

> Pragmatism begins … with phenomena concerning the *use* of certain terms and concepts, rather than with things or properties of a non-linguistic nature. It begins with linguistic behavior, and asks broadly anthropological questions: How are we to understand the roles and functions of the behavior in question, in the lives of the creatures concerned? What is its practical significance? Whence its genealogy?   (Price 2011, 231–2)

> You will be a pragmatist about an area of discourse if you pose a Carnapian external question: how does it come about that we go in for this kind of discourse and thought? What is the explanation of this bit of our language game? … The explanation proceeds by talking in different terms of what is *done* by so talking. It offers a revelatory genealogy or anthropology or even a just-so story about how this mode of talking and thinking and practicing might come about, given in terms of the functions it serves.   (Blackburn 2013b, 75)

Three motifs characteristic of Cambridge pragmatism can be extracted from these passages. The first is Cambridge pragmatism's *agent-centredness*: in addressing philosophically puzzling topics, it does not start by asking about the *object* X, but instead centres on agents and their dispositions, attitudes, concepts, and words revolving around X.[27] The second motif concerns the type of explanation that is then sought: Cambridge pragmatism takes a *function-first approach* to the explanation of concepts or terms, relegating inquiries and explanations looking primarily at their content or meaning in favour of inquiries and explanations looking primarily at their point or function in practice. The third motif is the idea that pragmatism might be given a *genealogical* cast. Price and Blackburn say little about the distinction between synchronic and genealogical explanations of the functions of concepts. But when this third motif of a genealogical dimension is added to the preceding two, we arrive at the form of pragmatism that is at work in pragmatic genealogy:

$$\text{Pragmatic Genealogy} \quad = \quad \begin{array}{l} \text{Agent-Centredness} + \\ \text{Function-First Approach} + \\ \text{Genealogical Dimension} \end{array}$$

---

[27] Price's wording might suggest that he is narrowly concerned with linguistic as opposed to conceptual behaviour more broadly conceived; but this impression is quickly dispelled by his application of the method in Price (2011, 2017, 2013). As he says, both terms and concepts are fair game for the Cambridge pragmatist, as are any other regularities and patterns in the agent's behaviour (such as dispositions or attitudes) whose function might shed light on philosophically puzzling topics.

As Blackburn's reference to Carnapian external questions reminds us, there is an internal and an external way to hear the question of why we came to think as we do. We can take an *internal* point of view and answer the question from *within* a given conceptual practice, pointing out, for example, that we think in terms of justice or cruelty because we are suitably sensitive to the justice and cruelty in the world. From the internal perspective in which we reason *from* or *with* the concept of X, the world appears to us full of X in a way that makes it seem inevitable that we came to think in terms of X—we think in terms of justice or cruelty because both plainly abound. But we can also step back from a given concept and take an *external* view of it, contemplating the world as including ourselves and our conceptual practice. And once we shift to this external perspective and reflect *about* the concept of X, we may find it harder to identify objects in the world that license our thinking in these terms: what is it that drives us to think in terms of justice or cruelty, or indeed in terms of values, duties, rights, possibilities, and probabilities?

It is here, as an answer to a Carnapian *external* question, that pragmatic genealogy enters the picture.[28] Of course, Carnap originally drew his distinction between internal and external questions to cast aspersions on external questions. But he had in mind external questions concerning the *existence* of the things referred to in a given area of discourse—questions like 'Do numbers exist?' For Carnap, such questions were either trivial (if heard as questions within that discourse) or meaningless (if heard as questions external to that discourse). Yet this does not bar us from asking external questions in a pragmatic rather than ontological key. We can still ask *what the point is* of using a certain type of discourse, such as discourse about numbers, while granting that the answer 'Because numbers exist' belongs with the internal perspective on that discourse and does not appear satisfactory from the external perspective. What we are looking for is a reflectively stable account capable of putting *enough* explanatory weight on the function of having a certain type of discourse to avoid putting *any* explanatory weight on the existence of the entities referred to in that discourse. This is not to deny the existence of those entities. On the contrary—when we find nothing from the outside that conflicts with the view of things we take from the inside, the pragmatic external explanation makes us comfortable with our conceptual practices, allowing us to relax into saying that 'there are values, duties, rights, possibilities, chances', and 'many statements concerning them are true' (Blackburn 2017, 63). Hence, as Blackburn also observes, it is not what pragmatists end up saying that distinguishes their position, but rather their explanatory story about *how we got there*.

---

[28] Craig explicitly draws the connection to Carnap and sometimes uses the Carnapian label of 'practical explication' (1986). But as he notes, he uses it to refer to something considerably different from Carnap's method (1990, 8).

2.3.3 Non-Reductive Naturalism-cum-Pragmatism

It is a familiar worry that naturalistic explanations drag high-minded ideals through the mud, failing to do justice to lofty aspirations by reducing them to animal needs and urges. And it is an equally familiar worry that approaching human affairs from a pragmatist angle invites one to take an unduly austere view of them—what Nietzsche called 'that colour-blindness of the *Nützlichkeits-Menschen*' (*BGE*, §204),[29] which reduces human thought as far as possible to instrumental valuing and instrumental reasoning[30] or flattens human affairs by treating them as exhaustively explainable in terms of highly generic needs (as if a reader of Tolstoy insisted that every page of *War and Peace* was fully intelligible in terms of the needs of Pleistocene hominins).

What is needed is a naturalist-cum-pragmatist approach which is non-reductive—which does not reduce the higher to the lower, all thought to instrumental thought, or all needs to generic needs. When we inquire into the point of our concepts, what we want is some insight into their instrumental relation to 'other things that we know that we need and value, but an insight which does not reduce them to the merely instrumental' (Williams 2002, 90); we want to avoid 'going straight to our actual society with the apparatus of functional explanation' in a way that would 'distort our understanding of our own cultural situation, debar us from seeing what is peculiar to it as opposed to others, and lead us to a stupid reductionism' (Williams 2002, 35). Instead, we want to make sense of what seems unrelated to human needs *in terms* of its instrumental relation to such needs, but without getting involved in reduction. What we want, in other words, is *explanation without reduction*, and as Williams suggests, the aim of the genealogical method is precisely to help one get it.[31]

This is not to say that reverse-engineering must be reductive if it lacks a genealogical dimension; on a charitable reading, as we shall see in Chapter 3, ascribing a point to some actual conceptual practice of ours by highlighting its instrumental relation to some of our needs can in principle be done non-reductively. But pragmatic genealogy is better at assuaging fears about reductionism, because it starts by taking the reductionism worry entirely seriously, and then labours to show precisely why it is ill-founded.

If we say that the point of a venerable idea such as the intrinsic value of truth is to serve the need to share information, for example, the reductionism worry will

---

[29] Roughly, 'the utility people'.

[30] Horkheimer's (2004, ch. 1) critique of instrumental reason, to the effect that reason degenerates into irrationality if it overemphasizes instrumental concerns, proved particularly influential in this regard. Arendt also criticizes 'the *homo faber* mentality' (1958, 82) for which human activity is paradigmatically technical and reasoning paradigmatically instrumental.

[31] Williams (2002, 35, 90); see also Williams (2006f, 137) and Craig (2007, 197–200).

be that this leaves us only with instrumental motives focused on the instrumental value of need-satisfaction, which amounts to a denial that there is such a thing as the intrinsic value of truth. This reductive tendency generalizes to other ideas, the worry goes, for if we view high-minded ideas as tools serving practical needs, our default assumption should really be that our conceptual relations to the world are purely technical or instrumental, transparently and single-mindedly geared towards need-satisfaction—that we think only in terms of resources to be exploited, and all our ideas are indexed to our needs.

Rather than having to deny that they meant anything as radical as that, as synchronic reverse-engineers would, pragmatic genealogists can grant that if we seek to explain ideas in terms of their tool-like character, this is an instructive place to start. Indeed, pragmatic genealogists typically set out from a state of nature in which all the worst fears about reductionism are realized: one in which there really is no intrinsic value of truth, for instance, and people's concern with truth really is merely instrumental. But then the genealogists spend much of their energy working through what approaches moving straight to our actual concepts neglect, namely the *instrumental inadequacy* of thinking in such purely instrumental terms. By setting out from thoroughly instrumentalized versions of ideas and showing *why* these would be instrumentally inadequate, pragmatic genealogies vividly bring out just why something genuinely new and distinct would have to arise, and why it would shed the connection to needs at the level of its conceptual content—by becoming a concept of something pursued for its own sake rather than as a means to an end, for example. The genealogical dimension brings out the *instrumental reasons for the de-instrumentalization of thought*. This is why the method is particularly good at highlighting the non-reductive nature of its explanation: it shows just why the reductive picture will not work and why pragmatic considerations themselves demand that concepts outgrow pure instrumentality. Similarly, the genealogical perspective avoids the reduction involved in explaining everything in terms of generic needs, because while it invites us to start from maximally generic needs—i.e. the most generic needs we can find that still illuminate our object—it then invites us to add more socio-historically *local* needs to make sense not just of the form our conceptual practices *would* have if we were as rough-hewn as the creatures initially depicted in the state of nature, but of the form they actually have, now and around here.

The explicitly fictional state of nature is thus uniquely useful in keeping apart the austere and thoroughly instrumentalizing descriptions that a good naturalistic and pragmatist explanation *starts out* from and the richer and far less instrumentalizing description that such an explanation *arrives* at once it has factored in various complexities. As long as the abstractions at the beginning of the genealogy are not mistaken for depictions of hominin prehistory, pursuing naturalistic and pragmatist explanation in different stages distributed along a genealogical axis avoids reductively distorting our understanding of our actual situation. Genealogy

keeps austere fictions and the rich outgrowths of history in their places, thereby preserving our right to refer without scare quotes to truth, knowledge, and justice.[32]

2.4 Weakening and Strengthening Confidence

Finding out whether a conceptual practice answers to needs, and what needs these are, is unlikely to leave everything as it was before the explanation was given. This is particularly true when we work not just one step back to the needs to which a practice immediately answers, as in simple forms of conceptual reverse-engineering, but many steps back to the needs that these needs themselves derive from, as in pragmatic genealogies. In addition to their explanatory significance, pragmatic genealogies can therefore also possess normative significance, subverting or vindicating what they are about. Where pragmatic genealogies are destabilizing or subversive, however, this is typically not a result of reducing the higher to the lower or the intrinsically valuable to the merely instrumentally valuable, but of relating conceptual practices to concerns we either fail to share or fail to endorse. And because pragmatic genealogies—like most functional explanations—initially identify respects in which something is somehow worthwhile for someone, they tend in the first instance not to prove destabilizing or subversive at all, but rather *vindicatory*.

   The idea that genealogies can also be vindicatory is one that Williams in particular emphasizes, holding up Hume's genealogical explanation of justice as an example. One might 'accept Hume's account', Williams writes, 'and still give justice, its motivations and reasons for action, much the same respect as one did before one encountered the explanation—or perhaps more respect, if one had suspected that justice had to be a Platonically other-worldly idea if it was anything' (2002, 36). To determine whether and in what respects a genealogy is vindicatory or subversive, he adds in a note, we must ask 'whether a genealogical explanation of an outlook or set of values is such, when it comes to be understood, as to strengthen or weaken one's confidence in them' (2002, 283n19).

   Taking up Williams's pointer, I maintain that pragmatic genealogies can affect our attitudes towards our conceptual practices by weakening or strengthening our confidence in them. A pragmatic genealogy will be vindicatory (or subversive) to the extent that it strengthens (or weakens) confidence in a conceptual practice by showing us the point of living by that conceptual practice, either *tout court* (e.g. by showing us the point of valuing the truth) or in certain contexts (e.g. by showing us the point of valuing the truth in talking about history). As I shall understand it, *confidence* is what binds us to conceptual practices where the chains

---

[32] See Williams (2002, 35).

of reasons they articulate come to an end; it is the sense of indubitability with which we engage in conceptual practices, putting concepts to work and accepting the considerations that guide and flow from their application.

While pragmatic genealogies do not bear directly on the reasons for or against *tokens* of a conceptual practice—they are not justificatory of a particular application of a concept or a particular expression of a value or virtue—they can yield *reasons*, of a kind to which confidence is responsive, because in showing us whether and how certain types of conceptual practices respond to certain needs, pragmatic genealogies bear on whether we have reasons for or against *cultivating that type* of practice (I expand on this in Chapter 9). As I use the terms, then, pragmatic genealogies, though not justificatory in the sense of providing *reasons for concept application*, can nevertheless be vindicatory by providing *reasons for concept use*: reasons bearing on whether to organize one's life along the lines articulated by the concept at all rather than on whether to apply the concept in a particular case or to draw consequences following from its application. The difference is well illustrated by the trial of Oscar Wilde. Pressed by the cross-examiner to admit that a certain story was blasphemous, Wilde resisted. Yet he quite rightly did not dispute that the concept of blasphemy *applied* to the story. Instead, he found that given his needs and concerns, he had no reason to *use* the concept of blasphemy in the first place—'blasphemous', he remarked, 'is not a word of mine'.[33]

Moreover, relating concepts to needs not only gives us a sense of whether to cultivate a conceptual practice at all, but also of *when* it is pointfully applied. A pragmatic genealogy can 'colour-code' a practice according to how pointful each of its manifestations appears in light of our needs. Paradigmatically functional cases then become recognizable as such, as do cases where the practice becomes pointless or overreaches itself. As we shall see in Chapters 5 and 7, for example, the demand for truthfulness has its place, but a good thing is taken too far when that demand grows into the attitude of *fiat veritas pereat vita*—let truth prevail though life perish.

In the chapters that follow, my primary focus will lie on how genealogies can be vindicatory. This is partly because while critical genealogies have been the subject of much attention—recent book-length discussions of critical genealogy include Saar (2007), Koopman (2013), Sauer (2018), and Srinivasan (manuscript)— vindicatory genealogies have remained comparatively under-studied, no doubt on the presumption that they are boringly conservative.[34] And indeed, why should we care about vindication? This is itself an important methodological question if we are to understand the motivations of the pragmatic genealogists.

---

[33] See Hyde (1973, 107).
[34] Though see Joas (2011), who draws on the work of Ernst Troeltsch to advocate what he calls 'affirmative genealogy'.

One thing that gives vindicatory genealogies their significance is the sense that many of our concepts may *not* fare well under genealogical reflection. Our modern condition is importantly different from the Aristotelian ideal of the *phronimos* who lives confidently within a unified and tensionless outlook. We have a strong sense of our outlook's being one among many, of its being the product of a tumultuous and tension-laden history, and of there being no guarantee that the concepts we ended up with actually help us to live. The picture is one on which we are *already* disturbed, and find ourselves in need of reassurance—reassurance that we do not need what we do not have, and that what we do have is worth having. I hope to show that there are truthful naturalistic genealogies to be told about many of our ideas which, even if they still appear offensive to the most Platonic or metaphysical sensibilities, will vindicate those ideas in the eyes of many and show them to be stable under genealogical reflection.

Taking a closer look at vindicatory genealogies will also show that the standard classification of genealogies into 'vindicatory' and 'subversive' soon reaches its limits. To say that there are two kinds of genealogies, subversive and vindicatory ones, obfuscates the respects in which even vindicatory genealogies have a critical edge. To say that something can be given a vindicatory genealogy is not to say that it is vindicated *tout court*; the vindication is relative to certain needs, and if these are needs that we either do not share or do not want to see satisfied, a vindicatory genealogy will be subversive in our eyes. Moreover, even when a concept serves more or less everybody's needs in unproblematic fashion, grasping what these needs are enables us to discriminate between situations in which a concept earns its keep, and situations in which the concept is applied pointlessly or overreaches itself. Thus, even a pragmatic genealogy that is in the first instance vindicatory already contains grounds for critique, for in revealing how a concept derives its point from a particular set of circumstances, it also calls applications beyond those circumstances into question. By understanding the rightful place of something in our lives, we gain a sense of when a good thing is taken too far.

It is worth differentiating the idea that pragmatic genealogies will in the first instance prove vindicatory from the superficially similar but far more conservative idea that we find the *wisdom of the ages* crystallized in the concepts we inherit. This Wisdom-of-the-Ages thesis has been associated with ordinary language philosophers such as J. L. Austin and the later Wittgenstein.[35] Both philosophers can give the impression of operating under a strong presumption that the battery of distinctions we find in our conceptual practices will be fundamentally sound—sounder, at any rate, than the distinctions philosophers are likely to dream up in an afternoon.

---

[35] See Williams (1995d, 218; 1995g, 487; 2005h, 34–7; 2014f, 43–4). For representative passages, see Austin (1961, 130; 1962, 62–4) and Wittgenstein (2009, §124). See Queloz and Cueni (forthcoming) for a discussion of the conservatism charge in the context of Wittgenstein's later philosophy.

This rather optimistic view of our conceptual practices is just the opposite of that which animates vindicatory genealogies. Vindicatory genealogies draw part of their interest precisely from a strong *doubt* over whether our conceptual practices will stand up to genealogical scrutiny. If we trusted in the wisdom of the ages, there would be no need for genealogical reflection. It is therefore not at all odd that Williams, the pragmatic genealogist who offers the most explicitly vindicatory genealogy, should also be the one who most explicitly objects to this Wisdom-of-the-Ages thesis and the conservatism it encourages (2014f, 43–4).[36] Williams suspects the Wisdom-of-the-Ages thesis of being a '*myth*, a fanciful picture of the past designed to justify certain activities in the present' (2014f, 44). Far from being confident in the wisdom of our inherited concepts, he condemns the 'emptiness and cruel superficiality of everyday thought' (1995d, 218) and insists that we take seriously 'the idea not only that we are deceived, but that we are deceived by forces worth worrying about, such as . . . the effects of tradition' (1995d, 219). What genealogical inquiry aims to do is to look beyond the effects of tradition to *find out* whether and why our concepts merit confidence. Whether they do so is, on this picture, very much an open question, and it is because the question is open that there is a point to genealogical inquiry. It promises to put us in a position where we can more critically promote those forms of thought that *merit* our confidence and abandon those that do not. The mere fact that pragmatic genealogies will in the first instance be vindicatory in no way entails that they will prove vindicatory all things considered. A feminist critic like Catharine MacKinnon (1989, 1993) might well come to agree with Austin that 'our common stock of words embodies all the distinctions men have found worth drawing, and the connections they have found worth making' (1961, 130), but take this to speak against our common stock of words.

2.5 Responsible Conceptual Engineering

Even when a conceptual practice turns out to be problematic all things considered, genealogy in a vindicatory key can be just as important to the enterprise of improving our practices as genealogy in a critical key. We saw that reverse-engineering can guide and justify attempts to alter our concepts through conceptual

---

[36] Williams duly notes that the charge of linguistic conservatism is often countered 'by reminding us that the study of existing uses was indeed a "begin-all"—after the investigation, then possibly reform'; but he still takes this to encourage conservatism, because it 'amounts to the proposal that no revolution or even reform can be mounted without a thorough sociological investigation of the *ancien régime*; and this is a proposal which in other contexts is rightly regarded as amounting to the proposal that there be no reform' (2014f, 43–4). Moreover, it is not least to the fact that hasty reformers of the past did not show such restraint that we owe the wealth of distinctions laid up in language: 'In language, as in politics, the conservative runs into the fact that the old is only what used to be new' (2014f, 44). See Krishna (2014) for further discussion.

engineering. But pragmatic genealogy is particularly good at showing us how we should *not* go about altering our conceptual practices by alerting us to what we stand to lose. As we shall see in more detail in Chapter 3, genealogy provides a more comprehensive view than a non-genealogical form of reverse-engineering would of the variety of ways in which, often unbeknownst to us, our conceptual practices perform a multiplicity of functions.

This is a third benefit characteristic of pragmatic genealogy: it facilitates *responsible* conceptual engineering.[37] It is one thing to realize that our conceptual practices are problematic and need to change, and quite another to secure the kind of understanding required to change them responsibly. Conceptual engineering as envisaged by Herman Cappelen in *Fixing Language* (2018, 34), for example, is *defect-based*: it encourages us to proceed by listing generic defects in concepts or words—emptiness, incoherence, inconsistency, vagueness, or objectionable effects on society, cognition, or theorizing—before going on to fix them across our conceptual repertoire. Conceptual engineering is then conceived primarily as a matter of fixing our concepts or words.

Yet responsible conceptual engineering requires one to understand not just what needs fixing or what is *bad* about our concepts, but also what is *good* about them—the variety of things they do for us when they function well. As Amie Thomasson also emphasizes: 'before removing a piece of a car engine, lines in a software program, or an organ from the body, it is always a good idea to begin with reverse engineering' (2020, 447). If we are blind to the panoply of respects in which our concepts are worth having, we risk wreaking havoc by abandoning them without grasping what is lost by doing so. Such caveats are not the preserve of Oakeshottian conservatism; Isaiah Berlin likewise argued that if revolutions have tended to issue in something entirely different from what they intended, it was because they are always at risk of seeing only the tip of the iceberg and producing unanticipated consequences by stirring up the depths (1996, 28–32). Pragmatic genealogy supports responsible conceptual engineering by systematically encouraging us to achieve a comprehensive view of the multiple layers of functions laid up in our conceptual practices—not to focus only on the functions that are particularly salient to us because they are historically recent or because they are in the spotlight of some ethical or political theory, but to step further back and ask what older, more familiar, and hence perhaps less striking functions our conceptual practices also perform.

Even when a conceptual practice clearly has objectionable effects, identifying the vindicatory aspects of its genealogy can help us distinguish between a situation in which we should abandon the practice wholesale and one in which we should aim to preserve the benefits of the practice while mitigating its disadvantages.

---

[37] I am indebted to Damian Cueni for this framing. See Cueni (2020), where he advocates responsible conceptual engineering in international law.

If one accepts the upshot of Hume's genealogy as presented in Chapter 4, for example, namely that the distinction between mine and thine plays a crucial role in lowering the potential for conflict below the threshold required for social coexistence, then Rousseau, who was struck by the ills that elaborations of the institution of property brought in their wake, will appear sensible in recommending not that property be abolished, but that adjustments be made in other parts of our lives to mitigate its pernicious effects. Hume's vindicatory genealogy encourages us to think that Rousseau quite rightly does not turn into Marx.[38] There will doubtless also be cases where there is no straightforward way of remedying the downsides of a conceptual practice while preserving what is important and beneficial about it. But then the value of vindicatory genealogy will lie in making it clear that this is the bind we are in. Revisionary interventions or engineering efforts aiming to change what is dysfunctional should be guided and constrained by a grasp of what is functional. This is why responsible conceptual engineering requires conceptual reverse-engineering. Insights into what concepts we need and why we need them are useful not only for retrospective, explanatory purposes, but also for prospective, action-guiding purposes.

2.6 Genealogy's Place in the Methodological Landscape

Despite the manifold attractions of pragmatic genealogy as a method, I do not take it to be the One True Method of philosophy. It has its place, but this does not mean that it should displace all other methods—on the contrary, the different ways we have of approaching conceptual practices stand to profit from mutually informing each other. We just saw that pragmatic genealogy can inform conceptual engineering. But it also interacts fruitfully with less closely related approaches in philosophy. Hypotheses about the point of conceptual practices might be informed by and checked against experimental philosophy's findings about folk intuitions, for example.[39]

But some of the less obvious interactions between pragmatic genealogy and other approaches can be illustrated if we concentrate on approaches to concepts and organize them according to whether they focus first on the *intension*, the *extension*, or the *function* of a concept.[40]

Intension-first approaches are paradigmatically exemplified by conceptual analysis, which aims to provide an explicit intension to be measured against the

---

[38] I am grateful to Damian Cueni for this way of putting it.

[39] According to Turri (2016), for instance, the findings of experimental philosophy suggest that folk intuition regards knowledge transmission as the point of assertion.

[40] A similar organization is proposed by Gardiner (2015), who distinguishes intension-first, extension-first, linguistic, and teleological approaches. Hannon (2019) advocates a function-first epistemology.

intuitive extension in the hope that the intension will explain why the extension has the boundaries it has. Here the guiding question is a version of the Socratic 'What is X?' question, namely: 'What are the conditions that are individually necessary and jointly sufficient for something to fall under the concept of X?' If successful, such an analysis yields a list of conditions or proprieties of use under which the concept picks out all and only the cases that form part of the intuitive extension. This kind of approach thus focuses first on identifying the concept's intension in order then to take it as a basis for explaining the concept's extension.

Extension-first approaches, meanwhile, make the extension rather than the intension their primary focus and seek to make sense of why the intension is as it is on the basis of a prior grasp of the nature of the extension. Despite this reversal in explanatory direction, the guiding question remains Socratic: 'What is X? What are the properties or unifying principles that make it X?' The intension might be the way it is, for instance, because it tracks natural kinds: objectively unified substances or clusters of properties that tend to appear together from one sighting to the next as a result of homeostatic mechanisms, shared chemical structures, or reproductive chains—as is the case, respectively, for the concepts *mama*, *milk*, and *mouse*.[41] The intensions of such concepts are as they are because this renders concept-users suitably sensitive to homeostatic clusters of properties which precede the concepts. This strategy is not restricted to natural-kind terms, however. It is pursued more generally by attempts to understand our concepts via metaphysical inquiries into the nature of the things we think or speak about. Such inquiries begin with, say, mathematical, modal, or moral statements we take to be true, and then ask what *makes* them true, or what their terms *refer to.* Mysterious truth-makers such as numbers, possibilities, or moral facts then come under metaphysical scrutiny aiming to discover what numbers, possibilities, or moral facts really are.[42]

Function-first approaches, by contrast, focus primarily on the *function of having* a concept with such an intension and extension. While intension- and extension-first approaches are guided by versions of the Socratic Question, function-first approaches are guided by some form of the Pragmatic Question: 'Why do we have the concept of X? What does it *do* for us? What is the practical value of living by a concept that delineates just *this* extension by means of just *that* intension?' Once identified, the concept's function can then serve as a basis for explaining why the concept has the intension and extension it has. In this schema, pragmatic genealogy is a species of conceptual reverse-engineering, because unlike other approaches within that genus, it does not simply reverse-engineer the

---

[41] The examples are Millikan's (2000, 50). The conception of natural kinds as homeostatic clusters of properties is developed by Boyd (1991) and Kornblith (2002, 61; 2011).

[42] My fairly loose usage of the term 'metaphysical' follows that of Price (2011, 14), M. Williams (2013), and Blackburn (2013a). A more specific sense is given to it in Chapter 5.

concept's function synchronically, by looking at our present conceptual practice, but *via* the concept's genealogy; and conceptual reverse-engineering is in turn a species within the genus of function-first approaches, because it primarily seeks to identify the function of a concept in order thereby to explain why we have that concept (only a species, however, because there are also function-first approaches that do not reverse-engineer at all, either because they already know the function or because they focus on the function a concept *should* serve).

Intension-, extension-, and function-first approaches are not best understood as seeking to put each other out of business, and I certainly do not want to claim that the particular function-first approach that is pragmatic genealogy should simply replace all other methods. My point is rather that it is usefully incorporated into a pluralistic methodological repertoire in which one type of approach can be used to guide or pave the way for others. It is true that the three types of approach cannot be pursued simultaneously, since they bestow explanatory priority on the intension, extension, and function of concepts, respectively, and only one corner of that triangle can form the apex at any given time. Yet they are not mutually exclusive in the strong sense that would exclude their complementing, augmenting, or informing each other. As Peter Strawson insists, 'the kinds of concept we employ are not independent of the kinds of purpose for which we employ them' (1963, 506); nor, indeed, are the kinds of purpose for which we employ them independent of the kinds of concept we employ.

Before developing a pragmatic genealogy, therefore, one might profitably consult extant attempts to specify the intension of the concept at issue. Extant analyses of the concept may helpfully broaden or sharpen one's sense of the various properties that the concept might be tracking. The findings of extension-first approaches can similarly prove useful to pragmatic genealogy: understanding that a concept's extension is unified by certain natural principles might yield a clue as to what function the concept performs by giving one an independent grip on the projectability of its extension's properties from one sighting to the next.[43] And what we take ourselves to know about the intension or extension of a concept might not just inform the initial functional hypotheses that get the genealogy off the ground, but also offer an external validation of the genealogy once it is complete. To take the function of a concept as *explanans* is to treat its intension and extension as *explananda*. A good function-first approach to the concept of knowledge, for example, ought to be able to account for the intuitive appeal of the various definitions that have been proposed for that concept. If a pragmatic genealogy such as Craig's (see Chapter 6) can explain why competing conceptual analyses of the concept of knowledge have highlighted

---

[43] See Millikan (2000, 2017).

just the features they have, and why each has some intuitive appeal, then that will count strongly in the genealogy's favour.

Function-first approaches can in turn inform and guide attempts to pin down the intension or extension of a concept by specifying the *task* that a concept with the sought intension or extension must live up to: by understanding what needs a concept answers to, we gain a sense of what its intension can be expected to look like, and which aspects of the world it can be expected to lock on to. Understanding what is worth talking about given our needs can be a good guide to the analysis of our concepts. For example, Laura and François Schroeter (2015, 428–36) argue that the manifold needs that give our concept of *water* its point— not just in the chemistry laboratory, but also in the kitchen, in religious rituals, and in symbolic and artistic contexts—should inform our analysis of the concept. Similarly, Georgi Gardiner (2015, 40) points out that looking at the needs a concept serves can tell us how precise or fine-grained we should expect the concept to be: if the concept *adult* serves legal needs that can only be satisfied by a relatively precise and sharply defined concept, we should expect it to have a correspondingly precise definition—which it turns out to possess. If, by contrast, the concept *adolescent* primarily serves needs of stereotyping, marketing, and the media, vaguer boundaries will do. They might even do better. Sharp edges can be a defect: if you asked me for a bread knife and I gave you a razor blade because it was sharper, you would complain.[44] Realizing that our needs are best served by a concept with vague boundaries would give us reason to look for a definition that preserves this vagueness.

Last but not least, a function-first approach to a concept can also inform intension- or extension-first approaches *negatively*. It might suggest that we should not look for a conceptual analysis of that concept in terms of necessary or sufficient condition, or that a metaphysical inquiry into the nature of its extension is unlikely to reveal much of substance. Here also Craig's genealogy offers the paradigmatic example: his account of the concept of knowledge suggests that there are good practical reasons why the concept does not admit of a strict definition in terms of necessary conditions. In this case, pragmatic genealogy can rightly be said to *displace* conceptual analysis—it not only gives us reason to doubt that the search for a strict definition could succeed, but completely obviates the need for it. It only does so on a case-by-case basis, however. Craig's account may lead us to question the presumption that the search for strict definitions will be successful in other cases, but it does not by itself imply that we should eschew conceptual analysis across the board and resort to pragmatic genealogy instead— though by revealing philosophers' decade-long chase after a strict definition of knowledge to have been mere tail chasing, a pointless going around in circles due

---

[44] The example is Wittgenstein's (2000, MS 120, 142v).

to a failure to realize how the pieces they already had fit together, Craig can hardly avoid suggesting that conceptual analysis without pragmatic genealogy is blind.

Having motivated the project of conceptual reverse-engineering and situated it within the broader methodological landscape, it is time to take a closer look at the function-first approaches on offer within the genus of conceptual reverse-engineering, and in particular to compare pragmatic genealogy with a non-genealogical rival in order to understand the rationale for reverse-engineering through pragmatic genealogy.

3
When Genealogy Is Called For

A methodological choice faces those who decide to reverse-engineer the point of a conceptual practice: should they turn directly to its present manifestation and try to discern the point of the practice in the often overwhelmingly complex and history-laden tangle of our actual ways of going on? Or should they work towards the point of the practice indirectly, via a fictionalizing and historicizing genealogy that begins with the point of simpler forms of our actual practice? And are there cases for which pragmatic genealogy is uniquely suited, so that this choice of method could be grounded in the features of the practice in question?

It is tempting to think that the genealogical approach is never truly called for, and that a more direct approach is always at least as good if not better because of its greater methodological perspicuity. Indeed, that view seems conclusively vindicated now that Miranda Fricker, once herself a proponent of pragmatic genealogy (2007), has moved to a non-genealogical form of reverse-engineering which she describes 'as a more straightforward and transparent way of achieving the very same explanatory pay-off' (forthcoming, 4) that pragmatic genealogies purport to deliver.

What motivates Fricker's move towards a non-genealogical sibling of pragmatic genealogy is the hope of reaping its explanatory pay-off without having to resort to its singular mixture of historicizing and fictionalizing. This mixture has indeed proved off-putting, and many have failed to see much of a point in giving this kind of genealogy (E. Fricker 2015; Hannon 2019, 52–3). Even those sympathetic to genealogy have tended either to endorse the historicizing while seeing no point in the fictionalizing (Dutilh Novaes 2015; Hacking 2005, 168; Humeres 2018; Koopman 2009; 2013, 71; Papineau 2019; Rorty 2002), or to endorse the fictionalizing while seeing no real need for the historicizing (Blackburn 2013b; Craig 2007; Pettit 2018, ch. 1; Price 2011; Weinberg 2006; M. Williams 2013).

Hence Fricker's attempt to offer a way around these worries: she proposes an interpretation of pragmatic genealogy on which its genealogical dimension is really a metaphor for explanatory priority, so that pragmatic genealogy can be viewed as an extravagant way of describing and ordering our actual practice (Fricker 1998, 2010b, 2016b, forthcoming). We can then safely dispense with the genealogical dimension: we can focus directly on a paradigm case—on a real instance of our present practice which we are willing to regard as displaying its most basic point—and try to make sense of other forms of the practice as derivatives of the paradigm case. This is Fricker's method of *paradigm-based*

explanation. In her paradigm-based explanation of blame, for example, the paradigm case is *Communicative Blame*: A wrongs B and B tells A with feeling that A is at fault. The point of this, according to Fricker, is to align A's and B's moral sensibilities (2016b, 167). She then seeks to derive from this an understanding of other types of blame, such as self-blame or blame of absent third parties, by showing how they serve the same overarching point in different ways—all without fictionalizing or historicizing.

By dispensing with the genealogical dimension while claiming to deliver the very same pay-off, Fricker of course renders acute the question of what, beyond colour and vividness, genealogies can add, and whether they are ever preferable to paradigm-based explanations. Why should philosophers interested in identifying the points of our current conceptual practices take the detour through genealogical fictionalizing and historicizing if the same pay-off can be achieved more directly? If even former practitioners of the method are jumping ship, why bother with pragmatic genealogy? This is a crucial question for the method of pragmatic genealogy, and one that the pragmatic genealogists themselves have failed to address in sufficient depth.

Accordingly, my aim in this chapter is to show when and why, if we want to understand our conceptual practices in terms of their point, we have reason to use pragmatic genealogy as I propose to understand it. I begin by developing a taxonomy of four increasingly complex forms of reverse-engineering which allows us to make sense of genealogical fictionalizing and historicizing as genuine and well-motivated *elaborations* of paradigm-based explanation. I then argue that it would be a mistake to replace pragmatic genealogies with paradigm-based explanations across the board, because pragmatic genealogy is called for in two kinds of cases: when dealing with *self-effacingly functional* practices, paradigm-based explanation gets a grip, but misses important aspects that pragmatic genealogy is better suited to capturing; and when dealing with *historically inflected* practices that lack a paradigm case or an obvious connection to generic human needs, paradigm-based explanation fails to get a grip altogether; by achieving a grip even here and helping us achieve a comprehensive view of what our practices do for us, pragmatic genealogy proves a valuable addition to our methodological repertoire.

3.1 Fictionalizing and Historicizing

Pragmatic genealogy is best understood as a genuine elaboration rather than a circuitous presentation of paradigm-based explanation. It goes beyond paradigm-based explanation in two respects: (a) it constructs *models*, in particular fictional prototypes of our practices; and (b) it introduces a *dynamic dimension* to help us understand how one gets from these prototypes to the practices we actually have. These are the two senses in which pragmatic genealogy can rightly be said to

fictionalize and historicize. But what exactly do we mean by 'fictionalizing' and 'historicizing'?

The fictionalizing of pragmatic genealogy is nothing mysterious; as I suggested in Chapter 1, it is best interpreted as the fictionalizing of model-building, which is to say that pragmatic genealogies diverge from reality only in that they involve one or several forms of idealization: 'Aristotelian' idealization by abstraction, i.e. the stripping away of non-essential features (Cartwright 1989); 'Galilean' idealization by distortion, i.e. the operation with assumptions known to be false (McMullin 1985); or the mixture of both which is known as 'caricature' (Frigg and Hartmann 2017). But the beauty of models lies in their power to provide us with a better understanding of their target system despite and even because of these abstractions and distortions. Various ways of accounting for this power have been proposed.[1] One such account, presented in terms that are particularly useful for our purposes by Alexander Prescott-Couch (2017, manuscript-b), holds that models isolate the *dependence structure* of a target system from the noise in the system in order to facilitate understanding of and practical engagement with that system. As Prescott-Couch argues, models of natural phenomena paradigmatically aim to display *causal* dependences while cutting out causal interferences in order to facilitate prediction and manipulation; models of arguments or views, by contrast, paradigmatically aim to display *rational* dependences between different propositions while cutting out mistakes and confusions in order to help us understand, communicate, and deliberate with others.

The dynamic models of pragmatic genealogies, I suggest, paradigmatically aim to display *instrumental* dependences: the ways in which conceptual practices are instrumental to the satisfaction of concept-users' needs. Idealization is of value here already because considering simplified prototypes of our practices in an uncluttered state of nature helps us break through the veil of familiarity to gain a sense of the more generic human purposes they serve. Idealization also cuts out noise, such as infelicitous conditions in which practices are temporarily prevented from serving a point. And once we add more historically situated needs to the model to do justice to the complex history of which our practices are the product, idealization helps us maintain a good overview by displaying the successive layers of practical significance in a neat and organized manner. This in turn informs our practical engagement with our practices: it motivates their cultivation if and when they prove worth having, and their abandonment if and when they prove pointless or worse.

Pragmatic genealogies proceed in two steps, with the first involving idealization, the second de-idealization: first, render plausible a hypothesis about why creatures like us would go in for an idealized, prototypical model of the practice

---

[1] See Weisberg (2013) for an overview.

we actually have—the 'proto-practice'; second, explain how we got from the proto-practice to the practice we actually have—the 'target practice'. The first step involves coming up with a hypothesis about the original point of the proto-practice and constructing a model offering a perspicuous representation of the proto-practice's instrumental relation to the needs of creatures like us. The model does this by showing how, already at this level of abstraction, the interplay of certain needs in that schematic environment generates a basic problem to which the proto-practice forms a salient solution. While there is some feedback between the processes of hypothesis formation and model construction—playing around with the model can suggest hypotheses and narrow down the search space—the hypothesis about the original point can initially seem fairly arbitrary. But this is not a problem, because it is the task of the subsequent genealogy to retroactively vindicate the initial hypothesis by demonstrating its fruitfulness and predictive and explanatory power.[2] The proof of the pudding is in the eating. Optionally, the pragmatic genealogy can also outline the *mechanism* by which the proto-practice might arise without presupposing possession of the relevant concepts in the originators: it might show how the proto-practice can be expected to emerge quite naturally, without much foresight or understanding on the part of those involved.

The pragmatic genealogist must then identify in what respects, if any, the proto-practice still differs from the target practice we actually have, for it is reaching something like the practice we know that provides what in the theory of models is called the *external validation* of the model (Kusch 2013). The second part of the genealogy must thus explain how one gets from the proto-practice to the target practice. To this end, the model must be de-idealized in the direction of our cultural situation by (a) describing the proto-practice's *primary elaboration*, i.e. its development driven by the practical pressures internal to the model, such as the foreseeable problems that the original solution offered by the proto-practice will bring in its wake; and (b) describing the proto-practice's *secondary elaboration*, i.e. its development driven by increasingly socio-historically local needs and the new problems that come with them. Both the primary and the secondary elaboration can be additive rather than substitutive, so that as needs are added, the proto-practice develops new forms alongside the old.[3] This helps account for the internal diversity and the criss-crossing relations of family-resemblance in the practice. In Hume's felicitous phrase, it presents the practice as having been 'warp'd into as many different forms' (*T*, 3.2.5.14) as our needs require.

Insofar as the dynamic models of a pragmatic genealogy successively incorporate ever more socio-historically local needs, they move beyond the categorical

---

[2] Hannon (2019, ch. 5) compellingly defends Craig against the arbitrariness objection as raised by Rysiew (2012) and Gerken (2017, §9.2.b).
[3] See Kusch (2009b, 2013) for an account of Craig's genealogy along these lines.

divide between *hypothetical* and *historical* genealogy:[4] they constitute a hybrid form that is still clearly an idealized model rather than a description of actual history, but that also genuinely historicizes the target practice by exhibiting it as the product of a complex historical accumulation of needs. The model is a receptacle for the insights of more orthodox historiography. But the model displays these insights in a format that is tailored to the purposes of philosophy. History informs the dynamic models of pragmatic genealogy so that these can in turn inform philosophical reflection and evaluation.

Finally, the charge that genealogy adds nothing but vividness itself already grants something important, namely that by combining fictionalizing and historicizing to lend its model the form of a genealogical narrative, the method does indeed add vividness. Doubly so, in fact: because it is a *narrative*, and because it is a narrative about how something is *made* or *constructed*. It is a pedagogical platitude that we understand an idea more easily and achieve a deeper and firmer intuitive grasp of its implications when it is couched in terms of a narrative rather than a list of abstract propositions. That effect is further magnified when the narrative is genealogical—when it is substantially about how human beings get from a stage in which some cultural phenomenon is absent to a stage in which it is present.[5] Genealogies do not just offer diagnoses of what something now does or what its function is. They get there via the story of *how something is made*, and there is a long tradition in philosophy of maintaining that what we understand best is what we understand how to make.[6] Nietzsche revives a version of this thought when he claims that genealogy can make sense of what is hard to analyse and impossible to define, as does Craig when he suggests that a narrative of conceptual synthesis can succeed where conceptual analysis failed.[7]

With this understanding of fictionalizing and historicizing in place, we can articulate a typology distinguishing four types of reverse-engineering that brings out just how pragmatic genealogy forms an elaboration of paradigm-based explanation:

---

[4] The divide is particularly clearly articulated in Gardiner (2015).

[5] Williams (2002, 20) defines genealogy as being about the genesis of cultural as opposed to natural phenomena, and as we shall see in Chapter 4, Hume's genealogical explanations grow out of his distinction between artificial and natural virtues. Even if we resist the claim that genealogies necessarily concern cultural phenomena, it is certainly a typical feature of them.

[6] Hobbes takes that line, for example, to argue for our privileged epistemic relation to geometric figures and social conventions: it is only what we made that we can know a priori or demonstrably, but once we realize that 'we make the commonwealth ourselves', as Hobbes put it, it can be seen to follow that 'civil philosophy is demonstrable' (1839, VII, 184); see Pettit (2008, 19–22). Fichte tries to exploit this idea to develop a scientific method for philosophy—the method of construction (Schmid forthcoming). But perhaps the most influential expression of this idea is Vico's principle that *verum et factum convertuntur*—the true and the made are convertible; see Berlin (2013, 39–57).

[7] See Nietzsche (*GM*, II, §13) and Craig (1993), especially the chapter entitled: 'Analysis? No, thanks'.

(1) *paradigm-based explanation*: identify an actual paradigm case of conceptual practice X, hypothesize its point, identify the needs it answers to, and use this instrumental relation to certain needs to elucidate practice X. (1) is exemplified by Fricker's account of blame (2016b).

(2) *prototype-based explanation*: construct a model of target practice X, hypothesize the point of this proto-practice, identify the needs it answers to within the model, and use the model to identify analogous and dis-analogous instrumental relations in target practice X. (2) has the character of Wittgensteinian explanations in terms of fictional objects of comparison whose similarities and dissimilarities to our practices are meant to elucidate them.[8]

(3) *generic pragmatic genealogy*: on the basis of an initial hypothesis about the original point of target practice X, construct a dynamic model showing why creatures like us would go in for a prototypical version of the target practice by identifying maximally generic needs generating a basic problem to which the proto-practice forms a salient solution; then consider the proto-practice's elaboration in response to further generic needs anticipatable from within the model; the closer this brings us to some generic form of the target practice, the better the genealogist's claim to having identified its practical origins and what it does for us. (3) is exemplified by Craig's genealogy of the concept of knowledge (1990).

(4) *pragmatic genealogy tailored to a socio-historical situation*: on the basis of an initial hypothesis about the original point of target practice X, construct a dynamic model showing why creatures like us would go in for a prototypical version of the target practice by identifying maximally generic needs generating a basic problem to which the proto-practice forms a salient solution; consider the proto-practice's elaboration in response to further generic needs anticipatable from within the model; then incorporate into the model increasingly socio-historically local needs which history tells us arose, and consider the proto-practice's elaboration in response to these more local needs; the closer this brings us to our local form of the target practice, the better the genealogist's claim to having identified its practical origins and what it does for us now and around here. (4) is exemplified by Williams's genealogy of truthfulness (2002)

---

[8] 'Our clear and simple language-games are not preliminary studies for a future regimentation of language—as it were, first approximations, ignoring friction and air resistance. Rather, the language-games stand there as *objects of comparison* which, through similarities and dissimilarities, are meant to throw light on features of our language' (2009, §130). And also: 'It disperses the fog to study the phenomena of language in primitive kinds of application in which one can command a clear view of the aim and functioning of words' (2009, §5).

and by his construction of a political concept of freedom tailored to our needs (2005c).[9]

Using this typology, the guiding concern of this chapter can then be re-described as being to understand when and why the move from (1) to either (3) or (4) is called for.

The answer I develop in the remainder of this chapter is that pragmatic genealogy proves its worth in two kinds of cases: when practices exhibit *self-effacing functionality*, and when they are so *historically inflected* that they lack a paradigm case or an obvious connection to generic needs. I shall argue that in the first kind of case, paradigm-based explanation gets a grip, but pragmatic genealogy proves the more powerful tool because it can add some insights that paradigm-based explanation misses; and in the second kind of case, only pragmatic genealogy gets a grip, thus earning its keep in the methodological repertoire.

3.2 Self-Effacing Functionality

Getting at the point of things is a powerful way of making sense of them, as attested by the pervasive interest in functional accounts shown and encountered by a wide array of disciplines. Yet functional accounts can seem to be at odds with our understanding of the phenomena they claim to illuminate, and there is an equally pervasive suspicion that we distort phenomena by forcing them into a functional mould. The result is often a stand-off between those who are impressed by the functionality of our practices and those who would defend them against what they perceive as undue reductionism. In its most extreme form, this stand-off opposes the *pure functionalist*, who insists that a practice can and should be exhaustively understood in terms of its subservience to our needs, to the *anti-functionalist*, who insists that these functional considerations are neither here nor there—the motivations for engaging in the practice are non-instrumental, and once we have spelled out what these are, there is nothing more to be said.

In virtue of its *hypothetical* and *dynamic* elements that distinguish it from the *actualist* and *static* approach of paradigm-based explanation, pragmatic genealogy is able to achieve a better grip on an important type of practice that gives rise to this stand-off between pure functionalists and anti-functionalists: practices exhibiting self-effacing functionality.

---

[9] For a related taxonomy, see Gardiner (2015). Gardiner's 'practical explication' corresponds to my (1), while her 'hypothetical genealogical teleology' corresponds to my (3). The present taxonomy differs from hers in two respects: it adds (2) as an intermediate type of reverse-engineering; and it adds (4) as a kind of hybrid between historical and hypothetical genealogy, which, on Gardiner's taxonomy, are distinct enterprises.

What does it mean for functionality to be effaced? Let us approach this question by setting out from a practice whose functionality is not effaced, but in plain view: the practice of queuing. Functionality plainly comes in at three levels: (a) the practice is functional *from a social point of view* in that it is instrumental to the satisfaction of collective needs and interests—it solves a coordination problem arising from the serial distribution of goods, namely the problem of determining the order in which people are served; (b) its functionality is central to people's *motivation* in engaging in it—it is also functional *from the individual point of view* and this provides an instrumental reason to queue; and (c) its functionality contributes to *explaining* why people engage in it. As a result, the claim that queuing is a means of achieving peaceful coordination raises no eyebrows, because it seems adequate at every level.

Now suppose that new, non-instrumental motivations to queue arose: motivations to queue not as a means to an end, but just because that is the kind of action it is. If these motivations were sufficiently widely internalized and reproduced, they might come to sustain the practice on their own. The practice would still be functional, but its functionality would recede into the background. In such cases, there is functionality, but the functionality *just ain't in the head* (to use Hilary Putnam's phrase). Participants in the practice are only dimly aware of it, if at all. This is what it is for functionality to be *effaced*—a term that helpfully ranges from being visible, but not as the primary motivation, to having dropped out of sight altogether. To understand *self*-effacing functionality, it then only remains to distinguish it from *contingently* effaced and from *necessarily* effaced functionality.

Where functionality is *contingently effaced*, participants are not primarily motivated by awareness of functionality, but contingently so—it is an accident of history that new motivations for engaging in the practice now loom larger in the participants' minds. If the participants were to be reminded of what the practice does, awareness of the practice's functionality would be fully compatible with confident engagement in the practice. Think of locally functional eating habits that adventitiously came to be primarily sustained by religious motives while losing nothing of their functionality.

Where functionality is *necessarily effaced*, participants are not primarily motivated by awareness of the functionality, and such awareness is radically incompatible with confident engagement in the practice: once the participants realize what the practice does, their commitment to it is instantly undermined. Think of the cases familiar from Critical Theory, where a practice only benefits a narrow class of people while being detrimental to the majority.

Where functionality is *self-effacing*, it is a functional requirement on the practice's functionality that participants not be primarily motivated by awareness of that functionality, but when they acquire awareness of it, this awareness is fully compatible with—and may indeed encourage—confident engagement in the practice on any reasonable conception of it. This last qualification is required

because one's conception of what the practice must look like to merit confident engagement may be so demanding as to exclude its performing such mundane offices as the satisfaction of human needs if it is to be worthy of respect. (Then there will be a tension between one's confident engagement in the practice and awareness of its functionality; but this may be taken to speak against holding so demanding a conception rather than against engaging in the practice.) A self-effacingly functional practice is thus functional, but only insofar as it is sustained by motives that are autonomous, i.e. not conditional on the practice's functionality in any given case. The practice must outrun its functionality in order to be functional. When this condition is met, the functionality of the practice will tend to show up, if at all, only as a secondary consideration among the contents of the participants' deliberation. But this, crucially, will be so for purely functional reasons, and not because the function is objectionable to the participants. The functionality is not just *effaced*, but *self*-effacing.

An example of a self-effacingly functional practice is *loyalty* to a group or cause: it is functional (let us assume) in that it stabilizes cooperative behaviour in ways that ultimately benefit most participants; but our motives in being loyal are not conditional on its fulfilling that function, and it is only because they are unconditional in this way that loyalty can fulfil its function. If loyalty is understood in purely instrumental terms, it will be mere window-dressing where it aligns with individual interest, and irrelevant where it does not: where the interests of the individual anyway align with the interests of the group, loyalty is functionally redundant; where they diverge, loyalty thus understood will not bring the individual to forfeit personal gain for the sake of the group and pull his or her weight in the cooperative venture; free-riding will be more attractive. And since the same is true for every individual in the group, the benefits of loyalty will be lost altogether. Hence, loyalty is unable to make a useful difference so long as it is understood in purely instrumental terms. To be functional at all, it must be more than just functional.

Self-effacing functionality is interestingly different from the two other forms of effaced functionality. It is not *contingently* effaced functionality, because there are good functional *reasons* why the functionality is effaced. At the same time, it is not *necessarily* effaced either, because the functionality can fully come back into view upon reflection without destabilizing the practice. Keeping one eye cocked on the functionality of the practice may even help sustain it. And if our confidence in a practice is wavering, highlighting the practice's functionality does not further weaken our confidence, but rather strengthens it by showing it to be *reasonable* confidence as opposed to the mere confidence of bigotry or fetishism.[10]

---

[10] The contrast between reasonable confidence and the confidence of bigotry is drawn in passing by Williams (2010). But see Fricker (2000), Hall (2014), and Queloz and Cueni (forthcoming) for further discussion.

More precisely, a practice exhibits self-effacing functionality if it meets the following conditions:

(1) *Functionality*: the practice is functional, i.e. it makes a useful difference to the lives of those who engage in it.
(2) *Autonomy*: the practice is sustained by motives that are not conditional on its functionality in a given case.
(3) *Dependence*: the practice can be functional only insofar as it satisfies (2), i.e. it would be unstable, redundant, or otherwise ineffective if sustained merely by motives conditional on functionality in a given case.
(4) *Explanatory connection*: the practice fulfils (2) because of (3), i.e. there is an explanatory connection between its autonomy and its dependence on autonomy.

We can call this special structure that is constitutive of self-effacing functionality the FADE structure (Functionality, Autonomy, Dependence, and the Explanatory connection between the latter two). In a practice exhibiting FADE, functionality will indeed fade from view. It will either be overshadowed by autonomous motives carrying more authority than instrumental considerations or it will be completely absent from the participants' consciousness. The FADE structure explains why this is so.

Why is pragmatic genealogy more adept at handling this kind of structure than paradigm-based explanation? After all, paradigm-based explanation *can* get a grip on self-effacingly functional practices. It is no objection to paradigm-based explanation that it is—in Strawson's (1974, 25) phrase—*one-eyed* in its focus on the practice's instrumental aspects. Sometimes we need to close one eye in order to achieve focus—to look beyond motivations to the overall point of the practice. This does not bar us from acknowledging that the participants' motives for engaging in the practice are of a non-instrumental nature.

But because paradigm-based explanation limits itself to taking an actualist and static view of our practice, it is more likely to miss the *functional connection* between the instrumental and the non-instrumental aspects of the practice, which explains *why* functionality is effaced. Paradigm-based explanation can discern functionality in a practice in which the participants are not motivated by functionality, and it can ask whether this is because awareness of the practice's function would be incompatible with engagement in it; but where the function is not obviously objectionable, the method does not by itself lead one to distinguish self-effacing from contingently effaced functionality.

Indeed, if we are very strict about the idea that paradigm-based explanation is concerned only with the actual, it cannot draw that distinction. The pressures on functionality to efface itself cannot be observed from actual practices alone. It takes some counterfactual thinking to see that *if* the practice were *not* sustained by

non-instrumental motives, the practice *would* be unstable. And it is only once we see this that we understand the *connection* between the functional and non-functional aspects of the practice: that we cannot abandon the non-functional aspects of the practice without also abandoning its functional aspects. Recognizing this involves contemplating a hypothetical, purely instrumentally motivated prototype and understanding the functional requirements on it to become non-instrumentally motivated.

By contrast, pragmatic genealogy allows us to understand *why* the functionality of the practice is effaced. We can construct a model of the target practice in which the functionality of the practice is entirely transparent—not only to us, the consumers of the model, but also to the agents *in* the model—and acts as the primary motivation to engage in the practice. By considering such a model, we can determine whether purely instrumental motivations would suffice to stabilize the practice. If not, this indicates that we are dealing with a self-effacingly functional practice, and that the *pure* functionality of the model is necessarily fictional: in reality, non-instrumental motivations are required to sustain the practice.

Pragmatic genealogy thus enables us to grasp why an actual configuration is as it is by constructing a fictional counterpart lacking some feature and seeing why this feature would be bound to develop. Even when it does not describe how the practice actually developed,[11] it can show that the non-instrumental understanding of the practice by the participants is a *counterfactually robust* feature of the practice. These modal insights into how it is impossible, possible, or necessary for creatures with certain needs to live yield dividends on two fronts. On the one hand, they have implications concerning whether it would be possible to pare away the non-instrumental aspects of a functional practice once we have identified its point. If the non-instrumental motivations are functional responses to the instability of the purely functional proto-practice, it would be misguided to eliminate them. On the other hand, the modal insights provided by pragmatic genealogy can bolster our sense of a practice's necessity and thereby defend the way we go on against what Amia Srinivasan (2011, 2019, manuscript) calls *genealogical anxiety*: the worry that the origins of our practices will turn out to be a source of discredit, or the more general worry that they will prove rationally contingent in being the result of forces that fail to vindicate them against possible alternatives, and hence fail to provide reasons why we should prefer this way of going on over possible rivals.

In sum, pragmatic genealogy deals fairly even-handedly with the stand-off between pure functionalists and anti-functionalists. Its treatment of self-effacing functionality allows function-first approaches to get a grip even in areas that are not obviously receptive to them: although self-effacingly functional practices

---

[11] See Queloz (2020) and Pettit (1996, 2000) for further discussion of how functional explanations can be informative even when they do not explain how something actually came about.

include non-instrumental aspects, *this fact itself* is shown to have a functionalist rationale: we reap the benefits of something like loyalty only if we are bloody-minded about it rather than benefit-minded.[12] But that same treatment also brings out the poverty of pure functionalism: an account that assumed that loyalty was functional but insisted on understanding it only in terms of instrumental motivations would be at a loss to explain its functionality. For the same reason, it would be unwise to suggest that we pare away the non-instrumental aspects of a self-effacingly functional practice once we discern its point. To put a new gloss on an old quip: 'Of course pragmatism is true; the trouble is that it doesn't work'.[13]

3.3 Nietzsche's Challenge: Historical Inflection and Local Needs

We now turn to the second kind of case in which pragmatic genealogy proves a valuable addition to our repertoire. Here a methodological remark of Nietzsche's provides a useful entry-point. It is well known that Nietzsche takes a dim view of philosophers' historical sense. But in the *Genealogy*, he rebukes the 'English genealogists' specifically for thinking ahistorically in assuming that there is an *instrumental connection* between our practices and timeless human needs.[14] Both we and our practices change, and philosophers will be led astray if they ignore the history that lies between the 'Darwinian beast' and the 'modern milquetoast'.[15] This amounts to a challenge—call it *Nietzsche's challenge*—for reverse-engineering. We might try and reverse-engineer the point of practices in an experimental spirit, to see how far we get on the assumption that the connection obtains. But Nietzsche's challenge is that the connection may well *not* obtain, because we or our practices have changed, and then reverse-engineering becomes—in more than one sense—pointless.

By putting a spotlight on the possibility of historical change, Nietzsche's challenge throws two features of reverse-engineering into relief. As long as it is understood as operating only with generic needs, two conditions must be fulfilled for reverse-engineering to work:

(i) The practice at issue must bear some instrumental relation to certain generic human needs—call this the *Generic Needs Condition*;

(ii) A paradigm case of the practice must be available that exhibits this relation—call this the *Paradigm Case Condition*.

---

[12] A phrase coined by Williams (2002, 59).

[13] The quip is attributed to Sidney Morgenbesser (Williams 2002, 285n14).

[14] Nietzsche (*GM*, P, §4, I, §2, II, §§12–13). Nietzsche's relation to 'English' genealogy is a topic of Chapter 5.

[15] Nietzsche (*GM*, P, §7).

Nietzsche's challenge puts pressure on both conditions, since a great deal of change at the level either of our practices or of our needs may well result in practices fulfilling neither the *Generic Needs Condition* nor the *Paradigm Case Condition*. Consequently, my aim in this section is to show how approaches seeking to reverse-engineer the functions of our conceptual practices can deal with Nietzsche's challenge, and how pragmatic genealogy in particular helps us do so. I shall first discuss the case in which the *Generic Needs Condition* is fulfilled while the *Paradigm Case Condition* is not; then the case in which it is uncertain even whether the *Generic Needs Condition* is fulfilled; and, lastly, the case in which neither condition is fulfilled.

3.3.1 Constructing Paradigm Cases

Consider the case in which the *Generic Needs Condition* is fulfilled while the *Paradigm Case Condition* is not: a practice still bears some instrumental relation to generic human needs, but it lacks a paradigm case exhibiting that relation. The problem for reverse-engineering is then not that the connection between the practice and generic needs is severed by historical change. The *Generic Needs Condition* still holds. But the function once discharged by a single practice may now be jointly discharged by a constellation of different practices into which the original practice has differentiated or by which it was replaced.[16] Or the problem may be that the *Generic Needs Condition* holds all too well—such a multiplicity of functions may have been layered into a practice that a paradigm case becomes difficult to identify. Where repurposing only 'obscures' previous functions, as Nietzsche puts it (*GM*, II, §12), practices can accumulate a rich historical deposit. There may then not be such a thing as *the* current point of a practice, because repeated alteration and repurposing have layered such a multitude of functions into the practice that neither a paradigm case nor an overall point can be recovered from the resulting mess. A good example—Nietzsche's own—is the practice of punishment:

> the previous history of punishment in general, the history of its exploitation for the most diverse purposes, finally crystallizes into a kind of unity that is difficult to dissolve, difficult to analyze and—one must emphasize—is completely and utterly *undefinable*. (Today it is impossible to say for sure *why* we actually punish: all concepts in which an entire process is semiotically summarized elude definition; only that which has no history is definable.) (*GM*, II, §13)

---

[16] A possibility that Williams pointed out in an interview about *Truth and Truthfulness*, where he noted that virtues can not only be reinforced but also replaced by institutions (Baker 2002). The function originally discharged by some virtue might then still be discharged, but the virtue itself would no longer be needed.

If we look back to the various functions a practice has discharged, we can discern 'how the elements of the synthesis change their valence and rearrange themselves accordingly, so that now this, now that element comes to the fore and dominates at the expense of the remaining ones, indeed in some cases one element (say the purpose of deterrence) seems to cancel out all the rest of the elements' (*GM*, II, §13). But when a practice's function only *seems* to cancel out previous functions, a complex and internally diverse deposit can form which defies not only analysis in terms of necessary and sufficient conditions, but also elucidation by means of a paradigm case exhibiting its core function. Nietzsche goes on to list eleven other functions layered up in the practice. 'Punishment', he concludes, 'is overladen with functions of all kinds' (*GM*, II, §14). In dealing with practices of this sort, paradigm-based explanation will either fail to get a grip on the practice or distort it by projecting a functional uniformity onto it which it does not possess.

Might one not try to capture the multiplicity of functions in a practice through multiple paradigm-based explanations? But where no single paradigm case stands out, identifying several only compounds the problem and must quickly become controversial and even arbitrary. Even if it can be done, it raises the question how the various functionality ascriptions relate to each other: which are competing hypotheses, and which are complementary? And how exactly do the complementary ones relate to each other? The concept of knowledge is another example of a conceptual practice that has invited multiple hypotheses concerning its function. Apart from Craig's contention that it serves to flag good informants and the numerous elaborations of that hypothesis (Greco and Hedden 2016; Hannon 2013, 2015, 2019; Henderson 2011; Kusch and McKenna 2018; McKenna 2014, 2015), its function has been thought to be to signal that inquiry is at an end (Kappel 2010; Kelp 2011; Rysiew 2012), to identify propositions we can treat as reasons for acting (McGrath 2015), to provide assurance (Lawlor 2013), to distinguish between blameless and blameworthy behaviour (Beebe 2012), or to honour the subject of knowledge attributions (Kusch 2009b).[17] What is missing is some kind of *master model* that indicates and integrates compatible functional hypotheses, giving us a controlled way of deciding which are basic and which are after-thoughts.

Does pragmatic genealogy fare any better? On the interpretation defended by Fricker, it does not.[18] What makes the two methods equivalent on Fricker's view is that she interprets pragmatic genealogies as not only starting out from a prototype, but as maintaining also that the prototype is really the core of our actual practice: 'The key is to see that . . . what is claimed about the State of Nature—for instance, that it contains a concept or practice with such and such features—is

---

[17] See Kusch and McKenna (2020, 1062), Gerken (2017, chs. 3 and 9), and the essays in Greco and Henderson (2015) for an overview.
[18] See Fricker (1998, 2010b, 2016b, forthcoming).

really a claim about what is basic (or "core") in our actual concept or practice'
(Fricker forthcoming, 7).[19] On this interpretation, the two methods stand on an
equal footing, because the *temporal* priority articulated by pragmatic genealogies
is a metaphor for *explanatory* priority within our actual practice. The time axis of
such genealogies serves as an expository device, organizing internally diverse
practices by isolating explanatorily basic features and successively bringing fur-
ther, increasingly complex but recognizably derivative features into view. This
interpretation licenses collapsing temporally extended genealogies into non-
temporal explanations as long as there is some other way of highlighting the
explanatorily basic—which is precisely the task shouldered by Fricker's paradigm
cases.

Fricker's interpretation combines two ideas to form what, in Chapter 1, we
called the *actualist* interpretation of pragmatic genealogy: (a) the narrative move-
ment in such a genealogy does not represent a temporal movement from our
conjectured hominin past to the present, and is therefore far removed from the
influential conception of genealogy as 'history, correctly practised'; (b) the primi-
tive form of a practice considered in the state of nature actually stands for a
paradigm case of our actual practice, and hence the genealogical derivation of the
less basic from the more basic can be safely collapsed into a description of our
actual practice.

This interpretation of pragmatic genealogy undeniably has much going for it.
By reading genealogy as involving neither an inference from fiction to reality nor
one from past to present, it alleviates worries about how fictional state-of-nature
stories can tell us anything about reality, and it deflects the charge of the genetic
fallacy, i.e. the alleged mistake of deducing claims about the present features of
something from claims about its genesis (a charge I return to in Chapter 9). The
main drawback of this actualist interpretation, however, is that it makes pragmatic
genealogy just as vulnerable to Nietzsche's challenge as paradigm-based explan-
ation: both, on this reading, move from one element in our practice that they
present as explanatorily and practically basic to other elements that they present
as derivative. This means that they both depend on there being, in our actual
practice, a paradigmatic core form which is functional given generic human needs
and which can give us an explanatory and critical grip on the practice. On the
actualist interpretation, paradigm-based explanation and pragmatic genealogy
stand and fall together.

But if, as I have suggested, we interpret pragmatic genealogy as *hypothetical* and
dynamic rather than as actualist and static, it *does* fare better than paradigm-based
explanation when paradigm cases are missing. We can join Fricker in maintaining
(a), that the narrative movement in a genealogy does not represent a temporal

---

[19] See also Fricker (2008, 47).

movement from our conjectured hominin past to the present, while denying (b), that it is a movement from paradigmatic to non-paradigmatic forms of our actual practices. Understanding genealogy instead as a movement from a strongly idealized model of a practice to a less idealized model of it—as a movement of de-idealization in the direction of our actual cultural situation—has the advantage that it does not depend on a paradigmatic form of the practice being extant: where history has failed to provide us with a paradigmatic form of the practice highlighting its functional relation to generic needs, pragmatic genealogy can *construct* one. We can organize and elucidate the complex amalgam that is our target practice using a simplified proto-practice. And we can then also model how we might have gotten here from there, where 'there' does not refer to some datable moment of emergence, but to an abstractly characterized basic predicament of which our present situation is a socio-historically local manifestation.

On this *dynamic model* interpretation, a pragmatic genealogy can serve as a master model that helps us *place* and *relate* the various further developments and acquired functions of the practice 'in a philosophical and historical space' (Williams 2005c, 76); not the three-dimensional space of our current practice, as the actualist interpretation has it, but the four-dimensional space along the quasi-historical time axis of the dynamic model. The genealogical model helps us situate, contextualize, and account for each of the different functions a practice acquired in different contexts, thereby imposing a form of order on the irreducibly varied synthesis that Nietzsche describes. The measure of the quality of that model will be its ability to make sense of the internal diversity of the practice and of the multiplicity of functions laid up in it.

Genealogy helps us identify, relate, and place the generic and local needs to which our conceptual practices answer. If one were specifically and *exclusively* interested in whether a practice serves some particular and highly local need, one would not need genealogy. But the pragmatic genealogist's ambition is to offer us a framework in which to situate and think about the dizzying variety of needs that have gone into shaping our concepts, and given this ambition, it makes sense to determine, in a peirastic or experimental spirit, whether we can illuminatingly start with generic predicaments that any creatures like us would be likely to face, even if we expect any such predicament to have in fact undergone a complex historical elaboration. Otherwise, we risk missing the most basic and invisible because all too familiar points served by some of our concepts. As Wittgenstein remarks, the 'aspects of things that are most important to us are hidden because of their simplicity and familiarity' (2009, §129). There are many important things that our ideas do for us which are buried deep within the historical deposit of our thinking. Pragmatic genealogies can help us excavate these old but no less important functions laid up in our ideas. That is where it is useful to ask: 'Of what is this thinking an enriched version?' (Heal 2017, 124). Contemplating a simple prototype and its point for creatures that are too indeterminate to feel

familiar can help us 'penetrate the blinding veil of ease and obviousness' that hides the workings of our own concepts from us.[20] Then, however, the pragmatic genealogist inquires into why these concepts were driven to *develop beyond* their most generic form by progressively adding ever more local needs into the picture. Adding these less generic practical pressures corresponds to enriching anthropological reflection with a posteriori socio-historical knowledge about more specific practical exigencies that certain concept-users came to face. We enrich our model of the most basic predicament to which a given concept answers with our knowledge of more local problems.

An important consequence of this interpretation of pragmatic genealogy is that making sense of a target practice using a proto-practice does not commit us to the further claim that the proto-practice is now extant as the *core* of the target practice. Craig (2007, 191), admittedly, seems to think both that his concept of proto-knowledge sheds light on the concept of knowledge *and* that it forms the core of our actual conceptual practice, but this has been deemed an unnecessary weakness in his account (Kusch 2011). Williams also sometimes uses the imagery of a core and its historical variations (2005c, 76; 2014g, 407), but whether he would be prepared to point to any actual instance of our practice of truthfulness as its core is less than clear; the core imagery seems to refer to what is central to his *explanation* rather than to the actual practice. The important point for our purposes, however, is that on the dynamic model interpretation, we can coherently maintain all of the following claims: first, that given certain highly generic needs that humans have anyway, they will need to see a certain function discharged, and we can illuminatingly construct a prototype of what a practice discharging it might look like; second, that this function is being discharged by our current practices; and third, that there is no one core form of our practice that directly corresponds to the prototype and conspicuously discharges that function. Just because some prototypical form of the practice is explanatorily basic does not mean that it is, or has ever been, extant.

The idiom of core and periphery or historical variation can be misleading in that regard, as it encourages thinking of the evolution of our conceptual practices on the model of a *snowball*: as the original core practice rolls down the slopes of history, it accumulates additional layers, but the core is still there by the time it ends up at our feet.[21] This is an improvement on the 'English genealogists' that Nietzsche rebukes for simply *equating* the current function of our practices with their original function. But we can take the injunction to think historically about functions a step further. Just as a snowball may encounter an obstacle that leads it to break up into pieces rolling down different paths and growing into different new shapes, a practice may differentiate into a *family of related practices* in

---

[20] The phrase is Austin's (1961, 128).
[21] The analogy hails from Kusch (2011, 19), who calls this the 'avalanche model' of genealogy.

response to a differentiation in the needs of concept-users. Between them, the practices may still *jointly* discharge the original function; but there may now be nothing that deserves being called the 'core'.

Whether or not such 'core' or paradigm cases are available cannot be determined a priori, and the value of paradigm-based explanation is best appreciated if we understand it as an attempt to find out: we hypothesize a candidate paradigm case in order to *see* whether a plausible paradigm case is available. But if not, I have been arguing, all is not lost for the reverse-engineering project. We can fall back on pragmatic genealogy and construct a prototype together with an account of why our actual practice differs from it.

3.3.2 The Roots of Continuity

So far, we have assumed that the *Generic Needs Condition* obtains: that the practices at issue still partly bear an instrumental relation to generic human needs. But why, once we are mindful of the wide array of contingencies and reinterpretations of which our practices are the product, should we remain confident that this is so? Nietzsche's challenge reminds us that we must be wary of the philosopher's foible of mistaking a failure of imagination for an insight into necessity.

Here also pragmatic genealogy proves a valuable addition to our toolkit. A pragmatic genealogy can be seen as an *argumentative chain* underscoring the assumption that a practice is a functional solution to some predicament we are bound to face on a continuous basis. It can reveal complicated instrumental relations between our practices and our needs even where we do not know they are there, and where the blank assertion that they obtain would not by itself be enough to convince.

The developmental narrative of a pragmatic genealogy can be thought of as a derivation of needs from needs: humans have a need for A, hence a need for $B, \ldots$, hence a need for X, where X is the prototypical form of the target practice. This gives pragmatic genealogy an edge over approaches that merely point out how a practice is functional relative to one particular set of needs—how blame serves a need for moral alignment, say, or how talk of probabilities serves a need to communicate and adjust our confidence in the occurrence of events. In the genealogical mode, we can present these fairly sophisticated needs as growing out of more primitive needs, and these out of even more primitive ones, until we reach what our audience is most likely to be willing to regard as uncontroversial needs—needs that are uncontroversially the genuine article (and not merely interests masquerading as needs) and that human beings uncontroversially have anyway. This enables us to derive needs we are *not* disposed to think we continuously have (e.g. a practical need to value accuracy and sincerity intrinsically) from

needs we *are* disposed to think we continuously have (e.g. a practical need for information and cooperation). When a function is being ascribed to something we did not necessarily expect to be functional at all, tracing out a chain of practical demands linking these high-flown abstracta to mundane concerns can make us more comfortable with the idea that these are functional responses to enduring challenges. Few will balk at the suggestion that our practice of thinking in terms of the concept *water* answers to enduring human needs; but the suggestion that concepts like *truth*, *knowledge*, or *justice* do the same may sound at first rather like the Panglossian claim that the bridge of the nose is there to rest glasses on. Pragmatic genealogy can help alleviate these worries by deriving needs we did not know we had from needs we knew we had, thereby revealing even the seemingly transient to be rooted in enduring human concerns. In this sense, the state of nature can act as a representation of the *roots of continuity* in the demands we face.[22]

Of course, the enduring presence of a need for something does not yet suffice to secure its emergence. Nor is it always sufficient for a single individual to realize that a practice would constitute a solution to a problem, for that solution may be inaccessible through individual instrumental reasoning—for instance, because it requires solving a coordination problem, or because it involves the essentially social process of constructing an intrinsic value.[23] But the genealogical perspective also allows us to show that the needed practice *could in fact have arisen* without implausible saltations. Genealogical explanation can add to our understanding by sketching mechanisms through which such obstacles might be overcome.

A pragmatic genealogy can thus do more to earn its conclusion than a non-genealogical ascription of functionality. If we accept, first, that the generic needs that operate as premises in the story are needs we share in some form; second, that the derivation of less primitive needs is valid; and third, that there are ways in which these might have driven us to develop certain practices in response, then we shall have been given a *reason* to expect there to be, in our actual cultural situation, some practice, or perhaps a constellation of practices, that is instrumental to the satisfaction of generic needs—and therefore a reason to think that the *Generic Needs Condition* obtains.

3.3.3 Incorporating Local Needs

The deepest engagement with Nietzsche's challenge, finally, is required when both the *Generic Needs Condition* and the *Paradigm Case Condition* cease to obtain. This will be the case to the extent to which practices are local outgrowths of

---

[22] I return to this line of thought in Chapter 9.   [23] Williams (2000, 156n6).

history that do not bear illuminating relations to generic human needs. Must reverse-engineering, and in particular pragmatic genealogy, lose all explanatory force once the connection to generic needs gives out, or can it still provide illumination even then?

It can—by augmenting our understanding of our practices, insofar as they fail to be amenable to elucidation by generic needs, with an understanding in terms of their point given *local* needs. We can ask what our practices do for *us* in particular, in contrast to other human beings that have lived.

It is tempting to think that local needs lie beyond the ken of reverse-engineering, because as a survey of the pragmatic genealogical tradition will show, pragmatic genealogists tend to make a point of grounding their explanations in highly generic and humanly basic needs. Does this imply a methodological restriction to universal or generic needs? Craig explicitly denies this:

> Any society that has a well-developed language … consists of creatures that have reached a considerable degree of mental complexity. Any number of different sorts of need may, for all we know to the contrary, follow in the wake of this complexity; so there is no a priori reason to think that we are tied by methodological principles to considering only needs of the very basic kind that I have actually tried to restrict myself to.   (1990, 4)

Understanding our conceptual practices as tools responding to our needs should not commit us to understanding them only in terms of generic needs. Such a focus on the needs we have anyway to the exclusion of needs we acquired or lost in the course of history would again be vulnerable to the Nietzschean charge of ahistorical thinking. It would make the method's explanatory basis unnecessarily narrow, thus either restricting the method's explanatory scope or rendering the method reductive if one insisted on explaining everything in terms of that narrow basis: to view all our conceptual practices as tools helping us cope with needs we have anyway would be to fall into the kind of reductive naturalism that assumes, as Robert Brandom (2011, 140) puts it, that we could safely dismiss Romantic poetry by asking what it has ever done for our biological fitness.

Rightly understood, pragmatic genealogies are not methodologically restricted to reverse-engineering the point of practices in light of universal or generic needs. They invite us to start off with the most general needs we can find that still bear an illuminating relation to the practice under investigation, but there is no reason why they cannot get a grip on socio-historically local practices by relating them to local needs. This is a strategy that is also open to paradigm-based explanation. But there are two respects in which the method of pragmatic genealogy proves particularly apt at dealing with local needs.

First, while paradigm-based explanation only relates current practices to current needs, pragmatic genealogy can exploit the fact that the connection between

needs and practices also holds dynamically: it additionally relates *changes* in conceptual practices to *changes* in needs. This equips it to answer Nietzsche's challenge by incorporating historical change into its dynamic model and explaining why ideas underwent further elaboration at certain junctures in history. Pragmatic genealogy thereby becomes able to explain both why we have an idea at all, in some form or other, and why we have it in the specific form it now has around here. Of course, *what* needs we find ourselves with will to some extent only be causally intelligible as the result of contingent historical change. But conceptual practices that answer to contingent needs are no less necessary for that. Insofar as the needs we contingently have generate real problems that necessitate solutions, the conceptual practices providing those solutions will be necessary for us. Given certain needs, however local, certain conceptual practices could not viably be different. They provide necessary solutions to contingent problems.

Second, pragmatic genealogy can offer what we might call a *comprehensive view* of a conceptual practice's relation to needs: one that brings out *both* the respects in which it serves generic needs and the respects in which it serves increasingly local needs while also *placing* and *relating* these aspects of the practice in its dynamic model, thereby situating them in a historical and philosophical space. It situates them in a historical space insofar as it shows which aspects of the practice are the product of highly general facts about us, and which are the product of more particular historical circumstances. And it situates them in a philosophical space insofar as it shows their relative importance and ineluctability. Do they answer to needs we can critically endorse? Are these needs we cannot but have, or are they needs we can eradicate by changing our circumstances? As Colin Koopman has pointed out in connection with Deweyan pragmatism and Foucauldian genealogy, pragmatist inquiry tends to uncritically take as given the problems to which it sees practices as answering; by thematizing the genesis of the problems themselves, genealogy can redress that deficit (2011, 537). The same is true of pragmatic genealogy: by tracing problems to the needs from which they arise and situating these needs along a genealogical dimension, we move from a merely technical stance bent on problem-solving to a more critical stance capable of assessing whether these are problems we should be trying to solve.

The value of such a comprehensive view is that it safeguards us from two ways in which our view of a conceptual practice can be unhelpfully simplistic: by understanding the practice exclusively as a response to generic needs when it also answers to local needs; or by understanding it exclusively as a response to local needs when it also answers to generic needs. Both kinds of simplifications should be avoided, because we ideally want to understand all the respects in which a practice answers to needs—for explanatory purposes, of course, but quite particularly also for critical or revisionary purposes. If we fixate on generic needs that make a practice seem well worth having, we run the risk of missing the respects in which its local form also serves local needs that may be

problematic, and to that extent give us reason to alter or abandon the practice. Conversely, when we find that a practice serves local needs we find problematic, and to that extent have reason to alter or abandon the practice, this insight had better go along with a grasp of the extent to which the practice also serves generic needs of a very basic and easily overlooked sort, because this will show us how we should not go about tampering with our conceptual practices by alerting us to what we stand to lose.

Pragmatic genealogy, then, does much to help us meet Nietzsche's challenge. First, whereas paradigm-based explanation depends on practices including a suitably paradigmatic core form that is conspicuously subservient to human needs, pragmatic genealogy can construct a proto-practice. Second, pragmatic genealogy need not blithely assume that generic human needs are still informative, but can offer an extensive argument for this which might also reveal instrumental relations between our practices and our needs that we did not know were there. And third, insofar as Nietzsche's challenge undermines attempts to elucidate our practices in light of generic needs, pragmatic genealogy can still get a grip by drawing on local needs. In light of Nietzsche's challenge, paradigm-based explanation thus turns out not to be enough, which means that the fictionalizing and historicizing of genealogy cannot entirely be dispensed with. We need pragmatic genealogy, and we need to conceive of it not as a baroque form of paradigm-based explanation, but as a genuine elaboration of it that expands the repertoire of the reverse-engineer with dynamic models capable of situating generic and local needs in a historical and philosophical space.

While this chapter has focused on how pragmatic genealogy can add something to Fricker's approach, I believe that many of the considerations advanced here go wider. I suggested in Chapter 2 that Cambridge pragmatists like Blackburn and Price sometimes talk of offering 'genealogies'—of explaining 'how it came about' that we go in for such things as moral and modal discourse. But they draw no very sharp distinction between giving a paradigm-based explanation and giving a pragmatic genealogy.[24] Consequently, they leave unanswered the question of when and why a Cambridge pragmatist should opt for a genealogical as opposed to a non-genealogical approach. If the considerations advanced in this chapter have any truth to them, there is much to encourage the thought that here, too, past explanatory successes might be retained and new ones achieved by enriching one's repertoire with the dynamic models of pragmatic genealogy. Pragmatic genealogy can help us explain why certain practices outgrow their own functionality, and it can help us achieve a grip even on practices that do not revolve around neatly functional core cases. This is but a small part of what we need to

---

[24] Blackburn (2002a, 103; 2013b, 75); Price (2011, 12, 17–19, 29, 231; but see 320 for a brief discussion of contingency).

be able to do if we are to make sense not just of generic humanoids wielding purely functional tools, but of ourselves.

Having achieved a clear, if abstract, understanding of the method and its rationale, we can now turn to the history of philosophy to flesh out that understanding with concrete examples of the method. Chapter 4 locates the roots of the tradition of pragmatic genealogy in the eighteenth century, and more particularly in David Hume's treatment of ideas as remedies to inconveniences.

4

Ideas as Remedies to Inconveniences

David Hume

In the methodological tradition of pragmatic genealogy, the salient ur-genealogy is not Nietzsche's vituperative attack on Christian morality, but David Hume's vindicatory account of the artificial virtue of justice. The Enlightenment was the heyday of genealogical inquiries into the origins of human capacities, ideas, and institutions, and a common denominator of these Enlightenment genealogies was a naturalistic concern 'to present human artifice as natural—or to explain how human beings have naturally crafted their cultural and material environments' (Lifschitz 2012, 5).[1] Hume fits this description, but the third volume of *A Treatise of Human Nature*, published in November 1740, stands out as presenting us with a genealogy that has as good a claim as any to count as the first fully fledged pragmatic genealogy.[2] Indeed, it holds an important place within the history of genealogy *tout court*: the intellectual historian István Hont, noting that Isaak Iselin's *Über die Geschichte der Menschheit* (1764) is usually 'credited with inventing the genre' (2015, 42), points out that Hume's genealogy of justice is clearly more deserving of the title (2015, 35).

My concern in this chapter is to show how reading Hume as a pragmatic genealogist sheds new light on his account of justice, setting it apart from the kind of conjecture about our distant past that Dugald Stewart, writing at the dusk of the Scottish Enlightenment, called 'conjectural history' (1858, 34). I argue that viewing Hume's thought through the methodological lens of pragmatic genealogy

---

[1] See Tuck (1979, 174), Lifschitz (2012), and Palmeri (2016).

[2] Pieces of genealogical reasoning can of course be found throughout the history of thought. Srinivasan (manuscript) offers a historical overview that starts with a fragment from Xenophanes; see also Angehrn (2007). Whether Hume should count as inaugurating the pragmatic tradition of genealogy is a matter of where one wants to draw the line between genuine instances of a tradition and mere anticipations or intimations of it. One noteworthy forerunner is Bernard Mandeville's *Fable of the Bees* (1988 [1714]). But Rowett (manuscript), for example, argues that Hume's genealogy of justice is anticipated already in Plato's so-called 'Great Speech' in the *Protagoras* as well as in the *Republic*; Griswold (2018, 39–40) also links the genealogies of Hume—along with Jean-Jacques Rousseau's *Second Discourse* (1977 [1755]) and Adam Smith's *Theory of Moral Sentiments* (2002 [1759])—to Plato, and moreover identifies a forerunner of Rousseau's genealogy in Lucretius's account of the emergence of society in Book V of *De rerum natura*. Tuck (1987, 111–12) suggests that a pragmatic concern to exhibit certain laws or practices as functionally necessary to social coexistence is a characteristic aim of the state-of-nature tradition from Hugo Grotius onwards.

brings three under-appreciated aspects of it into focus.[3] First, it shows that the ingenuity of Hume's approach in 'Of the origin of justice and property' is that he seeks to understand what property *is* not by looking for the natural relations that make it property, as Locke had done, but rather by understanding property in terms of the practical origins of certain *human attitudes*, namely those encapsulated in the virtue of 'justice'. This corrects the impression—easily given by the text—that Hume is primarily puzzled by how the virtue of justice might have arisen, but that his idiosyncratically narrow conception of justice as respect for property leads him to note in passing that the idea of property would arise on the heels of the virtue of justice. I suggest on the contrary that Hume's genealogy of justice is not just incidentally, but primarily and essentially an attempt to make sense of the naturalistically puzzling phenomenon of property. He proposes to understand the nature of property in terms of human attitudes towards it, and particularly in terms of the virtue of respecting the possessions of others *as* their property—which is what Hume means by 'justice':[4]

> Our property is nothing but those goods, whose constant possession is establish'd by the laws of society; that is, by the laws of justice. . . . A man's property is some object related to him. This relation is not natural, but moral, and founded on justice. 'Tis very preposterous, therefore, to imagine, that we can have any idea of property, without fully comprehending the nature of justice, and showing its origin in the artifice and contrivance of men. The origin of justice explains that of property.   (*T*, 3.2.2.11)

Hume here exemplifies the general strategy of pragmatic genealogy: instead of taking an object-centred approach and trying to identify the objective nature of property (as Locke did when he delineated how the laws of nature determine natural property rights), Hume takes an agent-centred approach, looking at human attitudes and dispositions expressive of concern with property. He then tries to make sense of why these attitudes and dispositions might have arisen by showing that they serve a function given certain human needs, thereby taking a function-first approach with a genealogical dimension. He shows that 'tho' the rules of justice be *artificial*, they are not arbitrary', which is to say they are grounded in *reasons*: there are practical exigencies that make the 'invention' of norms regulating the handling of external possessions not only 'obvious', but

---

[3] Hume scholarship has only recently come alive to the genealogical dimension of Hume's thought. Lottenbach (1996), Wiggins (2006, ch. 2), and Blackburn (2013b, 78) speak of Hume's 'genealogy of morals', Price of Hume's 'genealogy of modals' (2008), and Kail sees Hume as systematically pursuing a form of 'genealogical naturalism' (2016).

[4] Hume often employs 'justice' broadly synonymously with 'equity' and 'honesty', and he sometimes uses the term to include the keeping of promises as well (*EPM*, App. 3, 7)—understandably, since the emergence of promising turns out to be connected to that of property.

'absolutely necessary' (*T*, 3.2.1.19). 'What other reasons could writers ever give, why this must be *mine* and that *yours*', he defiantly asks in the second *Enquiry*, 'since uninstructed nature, surely, never made such a claim?' (*EPM*, 3.2.30). Like other artificial virtues, the virtue of justice is revealed to be what Rachel Cohon calls a *prosthetic character trait*, developed 'to remediate our natural disabilities' (2008, 3). Reading Hume this way also shows his subsequent genealogy of *promising* to grow directly out of his genealogy of justice. The emergence of property brings in its wake a new practical problem to which the institution of promising forms the solution. In both cases, Hume seeks to account for the emergence of puzzling ideas by showing how they function as remedies to inconveniences.

Second, Hume is acutely aware of the dangers of intellectualism and circularity that beset genealogical explanations: the adoption of a conceptual practice may have a point for a collective of concept-users, but the genealogy risks being implausibly demanding or even presupposing the concept it seeks to explain if it represents originators as going through a process of deliberation resulting in an intention to adopt the practice *because* it serves that point. That is why Hume takes pains to spell out a path through which the virtues of justice and fidelity to promises might have arisen *without* making great demands on the understanding, foresight, or public-mindedness of those involved.

Third, precisely because of his commitment to avoiding intellectualism, Hume does not rest content with identifying instrumental relations that in fact obtain between needs and conceptual practices. He seeks to identify not only the rationale underlying the emergence of conceptual practices, but also—at a level of detail that is unique among the genealogists discussed in this book—the string of *motives* and the *psychological mechanisms* that might have driven concept-users to adopt those practices and to value adherence to them as virtuous. But before reconstructing Hume's genealogy to bring out these three under-appreciated aspects, we need to ask what motivated Hume's turn to pragmatic genealogy.

4.1 Motivating Genealogy: Artificiality and the Circle Argument

If human nature included a motive to justice and that motive naturally elicited approbation, then the motivational story of how we came to go in for the virtue of justice would be trivial. But Hume views both the practice of justice and its status as a virtue as cultural achievements rather than natural givens. Echoing the ancient distinction between what is given by nature (*physei*) and what is posited by humans (*thesei*), he distinguishes *natural* from *artificial* virtues. Natural virtues are dispositions that we naturally feel moved to manifest and that naturally elicit the approbation of others, thanks to what human beings are like anyway. Artificial

virtues, by contrast, are dispositions towards which we have no straightforward natural motive, and which only 'produce pleasure and approbation by means of an artifice or contrivance, which arises from the circumstances and necessities of mankind' (*T*, 3.2.1.1).[5] Examples of artificial virtues include justice construed as respect for property, fidelity to promises, and allegiance to the government. Each of these is an artificial virtue 'in the sense that the attitudes and behavior that embody it are socially invented and this manufacture is then concealed, so that once the trait is widely internalized, it provides a prosthesis that lets human beings overcome their natural deficiencies' (Cohon 2008, 7). This means that artificial virtues invite, in a way in which natural virtues do not, the question of how and why they emerged—especially on a pre-Darwinian understanding of nature, where the contrast between the naturally given and the historically developed is still absolute.

Hume's genealogical inquiry receives additional impetus, moreover, from his argument to the effect that the salient present motive to justice—the sense of duty, rectitude, or morality—cannot explain how the virtue first arose. This sense of duty, which so impressed Immanuel Kant, is given a back-seat position in Hume's system.[6] While Kant subordinates virtue to duty, equating virtue with *fortitudo moralis*, the moral strength of will in fulfilling one's duty (Kant 1996, 164), Hume puts virtue before duty: the sense of duty may be a virtue, but it is a secondary one, which serves only to reinforce other virtues. It cannot play a foundational role among the virtuous motives.

Hume's argument for this conclusion, which has come to be known in the literature as the *Circle Argument*, is important for our purposes because it motivates Hume's genealogical inquiry. On Hume's view, actions must already *be* virtuous before we can come to feel that it is our duty to perform them, and the virtue of an action derives entirely from the virtue of the action's *motive* (*T*, 3.2.1.4). On pain of circularity, then, the sense of duty, rectitude, or morality cannot be invoked to explain the origin of justice.

This Circle Argument, from which Hume derives a number of requirements that motives must meet if they are to explain virtuous actions, can be set out as follows:

(P1)   The virtue of an action derives entirely from the virtue of the action's motive.

(P2)   An action must be virtuous before we can have a sense of its morality.

(C1)   The sense of an action's morality cannot be the *original* motive for the action, since that would be circular (though it does not exclude its later becoming a motive for the action).

---

[5] See Cohon (2006) and Kail (2016).   [6] See Blackburn (1998, 33).

(C2)   Therefore, for any virtuous action, there has to be some virtuous motive—other than the sense of its morality—from which the action derives its virtue (*T*, 3.2.1.4).

(C3)   Therefore, for any virtuous action, there has to be some motive in human nature that originally motivates the action (*T*, 3.2.1.7).

These requirements set the agenda for Hume's investigation of the virtue of justice. (C1) demands that the sense of morality be ruled out as *explanans*. (C2) demands that there be *some virtuous motive*, other than the sense of its morality, that first bestows merit on the just action. (C3) requires that there be *in human nature* some original motive, which need not be virtuous, to motivate its performance.[7] Hume's strategy is to identify a motive explaining why just actions first came to be performed in order then to explain the emergence of a different motive rendering just actions virtuous.

Turning first to the search for the 'original motive of justice' (*T*, 3.2.1.16), Hume considers three contenders: *self-interest* (*T*, 3.2.1.10), *public benevolence* (*T*, 3.2.1.11), and *private benevolence* (*T*, 3.2.1.13–16). None of them seem up to the task, since they either only strengthen our love of personal gain or are too weak to restrain it. Hume concludes that 'we have naturally no real or universal motive to observe the laws of equity [i.e. justice]' (*T*, 3.2.1.17). This passage has been read as suggesting that the quest for the original motive for justice fails (Gauthier 1992). But if we read the adverb 'naturally' as providing a crucial qualification here—a reading encouraged not only by the demand for coherence, but also by the fact that Hume added the word by hand to his first printed edition of the *Treatise*—it means only that no original motive of justice is to be found in our 'rude and more *natural* condition' (*T*, 3.2.1.9). What follows is *not* that there is no motive for justice apart from the sense of its morality, but that we should look for that motive outside our rude and natural condition. The argument then continues as follows:

(P3)   Any motive must be either a natural motive or an artificial motive.

(P4)   In the case of justice, we find no natural motive sufficient to explain the full range of just action.

(C4)   Therefore, there must be an artificial motive for justice that is distinct from the sense of its morality.

---

[7] (C2) and (C3) are often run together (Darwall 1993; Harrison 1981), and it is true that, in principle, one and the same motive may satisfy (C2) and (C3) by being both virtuous and originally found in human nature. Yet I follow Garrett (2007, 2015) in insisting that they be kept apart, for on the present reading, what first motivates just actions and what first renders them virtuous are indeed two different motives.

With the conclusion 'that the sense of justice and injustice is not derived from nature, but arises artificially, though necessarily, from education and human conventions' (*T*, 3.2.1.17), Hume's genealogical investigation has its work cut out: it must identify the *artificial* motive at the root of the virtue of justice and explain how it might have issued in something like the virtue we now have.

4.2 A Remedy to Conflict Over External Goods

What motive has 'both a sufficient force and a proper direction to counterbalance the love of gain, and render men fit members of society, by making them abstain from the possessions of others' (*T*, 3.2.2.13)? Since no other motive is a match for self-interest, Hume concludes that there 'is no passion, therefore, capable of controlling the interested affection, but the very affection itself, by an alteration of its direction' (*T*, 3.2.2.13). While Hume notes that self-interest in its natural state cannot constitute a motive for justice (*T*, 3.2.1.10), he hints that those defaults can be rectified by 'correcting and restraining the *natural* movements of that appetite' (*T*, 3.2.1.10). This is the possibility he goes on to exploit, arguing that self-interest can come to function as a motive for justice under certain conditions, namely when it is suitably restrained and corrected out of its natural shape by social forces.

For this socially modified motive of self-interest to come into play, the individual must first enter into society. As Hume points out, individuals have plenty of reasons to do so: society renders the individual 'in every respect more satisfy'd and happy, than 'tis possible for him, in his savage and solitary condition, ever to become', because society (i) augments power through the conjunction of force, (ii) increases ability through the division and specialization of labour, and (iii) reduces risk by fostering opportunities for mutual protection and reciprocity (*T*, 3.2.2.3).

Listing the advantages that the adoption of a new practice will bring is not sufficient to account for its emergence, however, since the individuals concerned may well be blind to those advantages. Hume recognizes this, granting that prior to experiencing them, human beings in their 'wild and uncultivated state' are unlikely to be 'sensible of [these] advantages' (*T*, 3.2.2.4). Nor can 'study and reflection alone' (*T*, 3.2.2.4) be counted on to make such advantages apparent to them. What Hume needs, therefore, are natural motives that will drive human beings into society even when they are blind to its advantages.

These natural motives are the 'appetite between the sexes' and the 'natural affection' (*T*, 3.2.2.4) of parents for their children. They naturally induce individuals to enter into a minimal social structure—what Hume calls a 'family-society' (*EPM*, 3.1.16)—that opens their eyes to some of the advantages that social structures can bring. And once they have experienced those advantages first-hand, human

beings will be in a position to purposefully strive to amplify those advantages by banding together on a larger scale.

Hume thus thinks that the 'passions of lust and natural affection' (*T*, 3.2.2.5) tend naturally to drive human beings into social structures. Yet he is also conscious of major threats to the stability of such structures. There are anti-social forces that are 'very incommodious, and are even contrary to the requisite conjunction' (*T*, 3.2.2.5), and it is as a product of these anti-social forces that Hume will seek to understand the need for the virtue of justice. His hypothesis is that the virtue of justice has its origin in the practical need to overcome a fundamental challenge to social stability (*T*, 3.2.2.12).[8] But what exactly is this challenge to which the virtue of justice forms the response? What disadvantages do humans suffer when they lack the virtue of justice?

Here a methodological difficulty comes into view: if the virtue of justice is as fundamental to social coexistence as Hume suggests, then, *ex hypothesi*, searching the historical record for a society in which the virtue of justice has not yet arisen will likely be in vain. If humans are bound to form societies, so that their 'very first state and situation may justly be esteemed social' (*T*, 3.2.2.14), and if the virtue of justice is necessary to life in society, then any society on record will already have overcome the problem we would like to explore.

For this reason, Hume points out, we may usefully extend our reasoning from the historical record, where social organization and justice are always already present, to a hypothetical 'state of nature' in which neither has yet arisen. Since Hume holds that in practice, it is 'utterly impossible for men to remain any considerable time in that savage condition', he conceives of the state of nature as 'a mere philosophical fiction, which never had, and never cou'd have any reality' (*T*, 3.2.2.14). Yet the 'idle fiction' of the state of nature nevertheless 'deserves our attention, because nothing can more evidently shew the origin of those virtues' (*T*, 3.2.2.14) associated with justice. The state of nature enables us to see that justice is

> intended as a remedy to some inconveniences, which proceed from the concurrence of certain *qualities* of the human mind with the *situation* of external objects. The qualities of the mind are *selfishness* and *limited generosity*: and the situation of external objects is their *easy change*, joined to their *scarcity* in comparison of the wants and desires of men.   (*T*, 3.2.2.16)

It is here, in presenting the virtue of justice as a remedy to inconveniences, that Hume makes the decisive move of transposing his genealogical inquiry into a pragmatic key. The state of nature sets out, in abstract and idealized form, a

---

[8] See also Hume (1978, 480; *EPM*, 3.2.17).

problem that most humans in most situations can be assumed to have faced in one form or another, and thereby allows us to see the target phenomenon—the virtue of justice—as a solution to that problem. The problem that Hume's state of nature lays out is that creatures like us in the kinds of environments we live in are likely to come into conflict over external goods. Human beings are naturally selfish, and whatever altruistic tendencies they may possess obey the inverse square law: they quickly peter out as one moves away from close acquaintances and kin. Whenever such agents find themselves in environments where resources are too scarce to satisfy everyone's needs and at risk of being lost to others, conflicts will break out over those resources. Communal life only exacerbates the temptation to prey on each other's possessions, a temptation that increases in proportion as society grows larger (T, 3.2.2.15). Given free rein, conflict over external goods must drive any society into violent chaos.

Hume's pivotal insight is that the community will be driven to tacitly coordinate on a practice that solves the problem of conflict over external goods without needing to get together and formulate promises or contracts—as Williams quipped, they can arrive at a convention without holding one (2014a, 109). All it takes is certain practical pressures combined with observation of others' behaviour, much as all it takes for rowers to row in unison is for each of them to possess the desire to advance and to pay attention to what the others are doing:

> Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to one another . . . it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. . . . In like manner are languages gradually established by human conventions without any promise. In like manner do gold and silver become the common measures of exchange.   (T, 3.2.2.10)

The emergence of the virtue of justice is conceived along similar lines: 'upon the least reflection', Hume writes, it becomes 'evident that the passion [i.e. self-interest] is much better satisfied by its restraint than by its liberty, and that, in preserving society, we make much greater advances in the acquiring possessions, than in the solitary and forlorn condition which must follow upon violence and an universal licence' (T, 3.2.2.13). Hume's insight is that if a sufficient number of individuals realize this, and if they realize that a sufficient number of individuals realize it, then they will coordinate on a set of conventions 'to bestow stability on the possession of those external goods' and to put 'these goods, as far as possible, on the same footing' with internal goods, i.e. the 'fix'd and constant advantages of mind and body' (T, 3.2.2.9). Out of this convention, the concept of *justice* as the *respect for this convention* arises, which in turn enables the establishment of the concept of *property*, which Hume explicates in terms of the

convention of justice as designating 'those goods, whose constant possession is establish'd by . . . the laws of justice' (*T*, 3.2.2.11).

Hume goes on to spell out these conventions governing the way possessions are treated as property (*T*, 3.2.3), and he ends up with four rules bearing a striking resemblance to the property laws of his own day: one's present possession counts as one's property (i) when one is the first to occupy and control it (the rule of 'occupation'), (ii) when one has long possessed it (the rule of 'prescription'), (iii) when it is the product of a resource that is already recognized to be one's property (the rule of 'accession'), or (iv) when it is consensually passed on to one from a parent or relation (the rule of 'succession'). Clearly, there is room for historical variation here, but the overall point of such conventions, whatever particular form they may take, is to remedy the problem of conflict over external goods by transforming possessions into property.

Reading Hume as offering a pragmatic genealogy that is gradually tailored to his own socio-historical situation allows us to see this not as an instance of a philosopher naively mistaking the peculiarities of his own time for anthropological constants, but as a perfectly legitimate de-idealization of a dynamic model in the direction of its target system. After all, the genealogy should make sense *of* the way property was understood in Hume's own day *to* the people who understood it that way. To be sure, his genealogy would have been stronger had he done more to *explain* why property norms came to have the specific shape they had in his day; but that he tailors his genealogical story to those property norms is entirely appropriate.

The heart of Hume's genealogical story is its explanation of how modified self-interest can power the emergence of a *new reason for action*—a reason to adopt and respect the conventions of justice. This emerging reason for action can be articulated as follows:

> *Instrumental Reason to Be Just*: Since we will all be better off as a result, it is in the interest of each of us to leave each other's possessions alone, *provided* others do the same.

Two things are worth noting about this reason for action. The first is that it is conditional in structure: it is in my interest to adopt the conventions of justice *if* others adopt them as well; otherwise, I would only harm myself by adopting them. This is why Hume remarks that it stands with justice much as it does with 'the building of a vault, where each individual stone would, of itself, fall to the ground' were it not supported 'by the mutual assistance and combination of its corresponding parts' (*EPM*, App. 3, 5). Second, though it is a reason for individual action, it is essentially possessed collectively; that is, its force as a reason derives from the fact that it is a reason for everyone, and that I can therefore expect others

to be moved by it as much as I am.[9] This helps overcome the challenge presented by its being conditional in structure. Hence, what drives the adoption of a set of rules ensuring the stability of external possessions is the

> general sense of common interest; which sense all the members of the society express to one another, and which induces them to regulate their conduct by certain rules. I observe, that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually expressed, and is known to both, it produces a suitable resolution and behaviour. And this may properly enough be called a convention or agreement betwixt us, though without the interposition of a promise; since the actions of each of us have a reference to those of the other, and are performed upon the supposition that something is to be performed on the other part.... It arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it.... This experience assures us still more, that the sense of interest has become common to all our fellows, and gives us a confidence of the future regularity of their conduct; and it is only on the expectation of this, that our moderation and abstinence are founded.   (*T*, 3.2.2.10)

Just as Hume appreciates that reasons alone will not propel solitary agents into society if those agents are unaware of them, he appreciates that convergence of interests alone will not suffice to give rise to the convention of justice. In addition, certain epistemic conditions must be fulfilled. In the case of entry into society, there was only one such condition: the reasons to prefer a social form of life must become visible to agents. In the case of justice, however, three epistemic conditions must be fulfilled. First, each individual must become sensitive to the fact that they have a reason to abstain from others' possessions, provided that others do the same—in Hume's terms, each individual must become 'sensible of an interest' in the regulation of his conduct. Second, each individual must realize that the same holds for others in the community—the sense of interest must be recognized to be common to all. Third, each individual must gain a sense that others recognize both these facts, for it is only then that the conditional preference (to abstain from others' possessions if others do) is paired with the expectation that its antecedent is fulfilled, and it is therefore only then that this preference is translated into action—in Hume's terms, the common sense of interest must be mutually expressed and known to all. In modern parlance, the first intimations of the virtue of justice arise when it becomes not just *shared* knowledge,

---

[9] A point highlighted by Williams (2002, 33).

but *common* knowledge that it is in each individual's interest to leave the possessions of others alone if others do the same—it is not just known to all, but commonly known to be known to all.[10] This, then, is the mechanism by which self-interest as channelled in the social context gets us from mere possessions to property.[11]

4.3 De-Instrumentalizing Justice

Prudence under social conditions does not yet make for virtue. If self-interest under social conditions satisfies the requirements (C1) and (C3) as the original motive for justice, it is hardly the virtuous motive bestowing virtue on just actions that would satisfy (C2). Adam Smith already complains that Hume occasionally sounds as if the utility of the virtue of justice were the main source of our approbation; and yet, as Smith caustically remarks,

> it seems impossible that the approbation of virtue should be a sentiment of the same kind with that by which we approve of a convenient and well-contrived building; or that we have no other reason for praising a man than that for which we commend a chest of drawers.   (2002, 4.2.4)

Our approbation may be 'enhanced and enlivened' by the realization that dispositions of justice have a point for us, Smith grants, but it is not our understanding of the 'utility or hurtfulness' of such dispositions that forms either 'the first or principal source of our approbation and disapprobation' (2002, 4.2.3). When we approve of an action, we are more concerned with the motive of the agent and the action's effects on particular individuals than with the long-term effects of this type of action on society. It is only 'men of reflection and speculation' (2002, 4.2.2) who would mistake such abstract considerations of social utility for the principal determinants of moral judgement. Smith takes Hume to be guilty of this philosopher's foible: 'The same ingenious and agreeable author who first explained why utility pleases, has been so struck with this view of things, as to resolve our whole approbation of virtue into a perception of this species of beauty which results from the appearance of utility' (2002, 4.2.3).

As Dennis Rasmussen has argued, however, Smith 'paints Hume as more of a moral utilitarian than he actually was' (Rasmussen 2017, 98). Hume can grant that we are not primarily motivated by considerations of utility—indeed, it is because

---

[10] The difference is crucial to many cooperative efforts, since it is often not enough if everyone knows that *p* (e.g. that the monarch should be deposed). As long as not everyone knows that everyone knows that *p*, no action will be taken—hence the importance of public rallies, manifestations, or demonstrations, which turn the fact that everyone thinks that *p* into common knowledge.

[11] See also Hume (*EPM*, App. 3.7; *T*, 3.2.2.10, 3.2.2.22).

he grants it that he feels the need to open his discussion of justice with the Circle Argument. But the upshot of the Circle Argument is that motivation by the sense of an action's morality will not do as the *original* source of our moral approbation of just acts. And if self-interest will not do either, Hume needs to explain how just acts came to be seen as virtuous acts.

There are, moreover, purely functional reasons why Hume cannot rest content with an explanation in terms of self-interest and considerations of utility: these are not sufficient to ensure the stability of the convention of justice. A convention is maintained only as long as all or most of the members of a society conform to it. But if too many people defect, this undermines others' expectations that the rules will be followed, which in turn means that it is no longer in their interest to conform either, and the convention breaks down. In the case of justice, non-conformity is encouraged by two things.

The first is that as population size increases, the 'shadow of the future' decreases—that is, the influence of anticipated future interactions on our current interactions wanes: 'on the first formation of society', Hume notes, self-interest is 'sufficiently strong and forcible' (*T*, 3.2.2.24) to maintain respect for property on its own. When a society is still small, the people one encounters are people one knows and interacts with daily. This means not only that our 'confined generosity' may go some way towards restraining our avidity, but also that a single injustice would soon become public knowledge and its consequences would be enduringly felt. But the larger a society gets, the more interactions tend to be mere one-time encounters between strangers. The threat constituted by a single act of injustice becomes ever less tangible to the individual, because 'when society has become numerous, and has encreas'd to a tribe or nation, this interest is more remote', and people do not as 'readily perceive that disorder and confusion follow upon every breach of these rules, as in a more narrow and contracted society' (*T*, 3.2.2.24). Consequently, in societies beyond a certain size, something more than self-interest is required to ensure the stability of the convention.

The second thing encouraging non-conformity is the temptation to free ride; that is, to profit from a collective good without contributing to it oneself, a temptation that arises whenever something benefits everyone while making it hard to exclude any individual from reaping those benefits. This threat is personi-fied by Hume's 'sensible knave'—the rational rogue who recognizes that the conventions of justice are in the general interest, but who feels no compunction about breaking the rules whenever 'an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy' (*EPM*, 9.2.9). 'Justice is the best policy' may be a good rule of thumb, the sensible knave reasons, but it is 'liable to many exceptions' (*EPM*, 9.2.9); the best policy may therefore involve being just *most of the time* while taking advantage of exceptions. As Hume grants, such reasoning is 'sens-ible'. So what is to prevent everyone from reasoning the same way and entraining

the demise of the convention? Hume must seek a first virtuous motive to justice not merely because justice now happens to be a virtue—the requirement articulated in (C2)—but because the explanatory merit of his entire genealogy depends on it: since the utility of the convention can explain why it arose, but not why it persists even in larger societies and in the face of free riders, a further motive is required.

Hume explains the emergence of the first virtuous motive to justice in terms of the psychological mechanism of sympathy. 'Sympathy', in Hume's usage, is the capacity to share the feelings and concerns of another. Like the sitar's 'sympathetic strings', which, though untouched, will resonate in response to vibrations of the primary strings (Baillie 2000, 52), sympathy enables us to resonate to the sentiments of others, even at the mere thought of them: 'the minds of men are mirrors to one another, not only because they reflect each other's emotions, but also because those rays of passions, sentiments and opinions may be often reverberated, and may decay away by insensible degrees' (*T*, 2.2.5.21). For Hume—as for his friend Smith—the capacity for sympathy forms the basis of our impartial concern for general welfare.[12] In the passages leading up to his genealogy of justice, Hume defends the thesis that moral approbation or disapprobation stem not from rational insight into abstract moral principles, but from *sentiments* of pleasure and displeasure. Combined with the mechanism of sympathy, this sentimentalist thesis promises to explain how 'reflecting on the tendency of characters and mental qualities' can be 'sufficient to give us the sentiments of approbation and blame' (*T*, 3.3.1.9): as we reflect on the tendencies of character traits and dispositions to affect others in positive or negative ways, sympathy with those affected produces pleasure or displeasure in us; and according to the sentimentalist thesis, these are the moral sentiments that find expression as moral approbation or disapprobation. For Hume, 'to have a sense of virtue, is nothing but to *feel* a satisfaction of a particular kind from the contemplation of character' (*T*, 3.1.3.3).

To account for the emergence of moral sentiments towards justice, Hume pursues the same explanatory strategy: when we observe or contemplate instances of injustice done to others, sympathy enables the harm incurred by *them* to generate displeasure in *us*. Likewise, the harm incurred by society as a whole also generates displeasure in us, and both forms of displeasure find expression as *moral disapprobation*:

> But though, in our own actions, we may frequently lose sight of that interest which we have in maintaining order, and may follow a lesser and more present interest, we never fail to observe the prejudice we receive, either mediately or immediately, from the injustice of others . . . when the injustice is so distant from

us as no way to affect our interest, it still displeases us; because we consider it as prejudicial to human society, and pernicious to every one that approaches the person guilty of it. We partake of their uneasiness by sympathy; and as every thing which gives uneasiness in human actions, upon the general survey, is called Vice, and whatever produces satisfaction, in the same manner, is denominated Virtue, this is the reason why the sense of moral good and evil follows upon justice and injustice.   (*T*, 3.2.2.24)

It is owing to the fact that 'sympathy interests us in the good of mankind' (*T*, 3.3.1.9) that we come to approve of motives that dispose people to respect the convention of justice. 'Thus *self-interest* is the original motive to the *establishment* of justice: but a *sympathy* with *public* interest is the source of the *moral* approbation, which attends that virtue' (*T*, 3.2.2.24).

While this moral sentiment towards justice arises from the contemplation of others, however, we 'fail not to extend it even to our own actions' (*T*, 3.2.2.24). This is because if the 'minds of men' truly are mirrors to one another, they reflect not only each other, but also each other's *reflections*. Sympathy can engender what one might call 'emotional ascent' (Blackburn 1998, 8): if sympathy makes it possible for the contemplation of others to produce displeasure in us, it likewise enables others' contemplation of *our* actions to produce displeasure in *them*, a displeasure that *we* may in turn come to share through sympathy. Hume offers a vivid example: 'A man will be mortified if you tell him he has a stinking breath; though it is evidently no annoyance to himself' (*T*, 3.3.1.26). Hence, it is because we 'naturally *sympathize* with others in the sentiments they entertain of us' (*T*, 3.2.2.24) that we extend our moral sentiments about them to our own behaviour. We begin to feel morally about our own disposition to respect the convention and develop a second motive towards justice, the sentiment that the rules of justice *should be respected, even by ourselves*. As a result, conformity or non-conformity with the rules of justice becomes in itself a reason for or against an action. This is a reason for action that is not derivative upon other reasons such as self-interest. Rather, a second, *independently motivating* reason to be just has emerged:

> *Independently Motivating Reason to Be Just*: In view of the displeasure I derive from infringements of the convention of justice by others as by myself, I disapprove of such infringements, and desire that the convention be respected by all, including myself.

This is the first virtuous motive required by (C2).[13] With the emergence of this motive, the disposition to justice is *de-instrumentalized*, i.e. it outgrows its purely

---

[13] *Pace* Baier, who holds that the first virtuous motive to justice is 'a sense of common interest', 'a special form of prudence' (2010, 66–7). By holding that it is a desire for justice generated by the capacity

instrumental subservience to self-interest. By becoming animated by a moral sentiment, the disposition to justice acquires an independent hold, and people can be moved to be just even in the absence of prudential reasons. Justice has developed a life of its own.

Once moral sentiments towards justice and injustice have arisen, they are stabilized and extended 'beyond their original bounds' (*T*, 3.2.2.25) by two amplifying mechanisms. The first is public praise and blame. Whenever societies reach a size where decision-making and administration cannot be handled solely through face-to-face discussions anymore, the need arises for decision-makers and political functionaries dedicated to maintaining the social order, and these will have a professional interest in upholding the conventions of justice. Hence, the sentiment of the virtue of justice is 'forwarded by the artifice of politicians', who, in order to 'govern men more easily, and preserve peace in human society, have endeavoured to produce an esteem for justice, and an abhorrence of injustice' (*T*, 3.3.1.25). The second mechanism consists in 'private education and instruction' to the same effect, as parents teach their children to 'regard the observance of those rules by which society is maintained, as worthy and honourable, and their violation as base and infamous', until those sentiments 'acquire such firmness and solidity, that they may fall little short of those principles which are the most essential to our natures, and the most deeply radicated in our internal constitution' (*T*, 3.3.1.26).

If these amplifying mechanisms are successful, they counteract the tendency of larger societies to lose sight of the fact that 'disorder and confusion' must result from breaching the rules of justice in the long run; and they enable societies to overcome the threat of the free rider, the selfish amoralist who is unmoved by the sense that a virtuous person should respect the rules of justice.[14] Philosophers have been wont to agonize over the question of what they can say to the amoralist who calls ethical considerations into doubt. They have sought an argument to show that the amoralist *must* recognize some ethical consideration on pain of inconsistency, an argument that will *force* the amoralist into the ethical life. Even if the amoralist could be shown to be subtly irrational, however, nothing forces him to be consistent (Nozick 1981, 408). The argument 'that will stop them in their tracks when they come to take you away' (Blackburn 2010, 294) is something of an *ignis fatuus* of moral philosophy.[15]

to see ourselves as others see us, I position myself between Garrett (2007, 2015), who takes it to be a disposition grounded in a desire to regulate one's actions by the rules of justice, and Cohon (1997), who emphasizes that the other-directed moral approval generated by sympathy itself becomes a motive to justice.

[14] Here I side with Baldwin (2004) against Darwall (1995, 317) and Gauthier (1992, 415–16), who think that Hume fails to address the problem that the free rider poses for his account.

[15] A point Blackburn takes from Williams (2011, 26).

Hume, by contrast, does not think we owe the sensible knave an argument. He grants from the first that the sensible knave is *sensible*, i.e. he cannot be faulted in terms of his *rationality*.[16] What Hume emphasizes is that he is *morally* at fault:

> I must confess, that, if a man think, that [the sensible knave's] reasoning much requires an answer, it will be a little difficult to find any, which will to him appear satisfactory and convincing. If his heart rebel not against such pernicious maxims, if he feel no reluctance to the thoughts of villainy or baseness, he has indeed lost a considerable motive to virtue; and we may expect, that his practice will be answerable to his speculation. But in all ingenuous natures, the antipathy to treachery and roguery is too strong to be counterbalanced by any views of profit or pecuniary advantage. (*EPM*, 9.2.10)

In Blackburn's apt phrase, 'what is wrong with the sensible knave is not that he has subtly trespassed against reason, but that he is a knave' (2008, 69). In the knave's case, the mechanisms of sympathy, private education, and public praise and blame have failed to engender the right moral sensibility for the virtuous motive to justice to get a hold. If we were all like him, the conventions of justice would not have been able to establish themselves in the way that they have. But most of us are not. The importance to us of 'inward peace of mind', 'consciousness of integrity', and a 'satisfactory review of our own conduct' (*EPM*, 9.2.10) ensures that we are not prepared to sacrifice the 'invaluable enjoyment of a character...for the acquisition of worthless toys and gewgaws' (*EPM*, 9.2.12). For us, the knavish life holds no temptation.[17]

4.4 Promising: Enabling Reciprocal Cooperation Over Time

It does not escape Hume's attention that even once possessions have been stabilized, people 'often reap but small advantage from it, while they are possess'd of a greater quantity of any species of goods than they have occasion for, and at the same time suffer by the want of others' (*T*, 3.2.5.8). Hence, the emergence of property in turn generates a salient need for *commerce* as the *transfer of possessions by consent*. But the problem is that not all goods allow for immediate delivery:

---

[16] Though Hume doubts that in practice, the sensible knave would succeed in implementing his knavish policy with profit (*EPM*, 9.2.11).

[17] Hume goes on to discuss the role of government as a check on rule-infringers and free riders: 'civil magistrates, kings and their ministers', he writes, 'being satisfied with their present condition, and with their part in society, have an immediate interest in every execution of justice, which is so necessary to the upholding of society' (*T*, 3.2.7.6). See Sagar (2018, ch. 1) for further discussion of Hume's account of the origin of government and how it builds on his genealogies of justice and promising.

One cannot transfer the property of a particular house, twenty leagues distant; because the consent cannot be attended with delivery, which is a requisite circumstance. Neither can one transfer the property of ten bushels of corn, or five hogsheads of wine, by the mere expression and consent; because these are only general terms, and have no direct relation to any particular heap of corn, or barrels of wine. Besides, the commerce of mankind is not confin'd to the barter of commodities, but may extend to services and actions, which we may exchange to our mutual interest and advantage. Your corn is ripe to-day; mine will be so to-morrow. 'Tis profitable for us both, that I shou'd labour with you to-day, and that you shou'd aid me to-morrow.   (*T*, 3.2.5.8)

The solution to the problem of how to transfer goods of this type is the institution of promising together with the virtue of fidelity to promises. Promises enable reciprocal cooperation over time—in particular, they make it possible to transfer goods and services that cannot be exchanged simultaneously. Hume's genealogy of promising starts with individuals recognizing that it is in their interest to exchange a greater range of goods and services than just those that permit immediate delivery. They recognize further that others have a similar interest, and hence that if they were to offer a good or service, they can reasonably expect it to be later repaid, because their cooperation partners are likely to reciprocate in anticipation of the potential benefits of future cooperation of that kind, which they can only reap if they do not now forfeit their status as trustworthy cooperation partners: 'I learn to do a service to another, without bearing him any real kindness; because I foresee, that he will return my service, in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me or with others' (*T*, 3.2.5.9).[18]

The emergence of this purely benefit-minded form of cooperation then generates the need to distinguish it from its less calculating lookalike, the 'more generous and noble intercourse of friendship and good offices' (*T*, 3.2.5.10). To answer this need and to mark the distinction between *interested* and *disinterested* commerce, a '*certain form of words*' is invented to express one's resolution to reciprocate in interested commerce. This form of words 'constitutes what we call a *promise*' (*T*, 3.2.5.10). As Hume makes clear, the '*symbols* or *signs*' (*T*, 3.2.5.10) in question need not even be verbal. The crucial thing is that they must possess a certain *practical significance*, namely that by using them, individuals *commit*

---

[18] It is crucial to this type of reasoning that the cooperation be perceived as open-ended: if a particular round of cooperation was known to be the last, the individual who went last would have no reason to reciprocate; anticipating this, the individual who went second-to-last would also be dissuaded from cooperating; and through backwards-induction, cooperation would fail to arise altogether—or so game theory predicts for the 'Centipede game' (where, in the original version, two players alternately get to decide whether to cooperate or defect over a hundred rounds with linearly increasing pay-offs); empirical studies suggest that people actually cooperate for some time before eventually defecting (Nagel and Tang 1998).

themselves to delivering on the promise, acknowledging the right of others to hold them to their promise, in case of defection, to sanction them by refusing future cooperation. The act of promising thus creates a new reason for action of the form: 'I ought to deliver on my promise because I do not want to lose my status as trustworthy cooperation partner'.

Much as the stability of possessions was secured by de-instrumentalizing justice, the institution of promising is buttressed by the fact that sympathy, together with private education and public praise and blame, transforms fidelity to promises into a virtue (*T*, 3.2.5.11–12). The result of this de-instrumentalization of the act of promising is that individuals feel motivated to keep their promises in part simply because they have promised, and people who are thus motivated are deemed virtuous.

Hume thus deploys a similar explanatory strategy in his genealogies of justice and promising. Its guiding insight is that practices that humans do not naturally feel moved to engage in can arise nonetheless when it becomes common knowledge that each individual stands to benefit from taking the initial step towards such a practice, because a limited amount of foresight and communication will reveal to each individual, first, that they have reason to make this initial step on the assumption that others cooperate; second, that they have reason to expect others to cooperate because the same reasoning holds for others; and third, that it is common knowledge that this is so. Together, these three considerations generate a new reason for action that is capable of driving the emergence of the beneficial practice, and the mechanisms of sympathy as well as private education and public praise and blame then add a second, moral reason for action by turning the disposition to participate in the new practice into a virtue.

But what exactly does the state-of-nature fiction do for Hume? And what broader lessons does his work hold for present-day pragmatic genealogists?

4.5 The Functions of the State of Nature

James Baillie remarks that Hume's account of the emergence of justice 'does not aspire to historical accuracy' (2000, 148). This is something of an understatement. Not only does Hume not even begin to situate his pioneers of justice and promising in space and time; he even pictures them in a 'suppos'd *state of nature*', all the while insisting that it is a mere fiction that never had and never could have any reality (*T*, 3.2.2.14–16). The fact that he openly advertises his state of nature as a counterfactual and counterpossible fiction suggests that he is not extrapolating from the known to the unknown to reconstruct, as plausibly as possible, a long-lost real history. *Pace* interpreters from Dugald Stewart (1858) to Simon Evnine (1993, 602), Hume is not offering conjectural histories—at least not in the sense in which these are usually interpreted, namely as an 'attempt to fashion a plausible

account of the earliest periods of human social life' in the absence of 'documentary or other material evidence' (Palmeri 2016, 1).[19] But if Hume does not aspire to historical accuracy, what other kind of accuracy does he aspire to?

Rousseau notes that some genealogical explanations are 'better suited to eluci-date the Nature of things than to show their genuine origins' (1977, 132), and thanks to the methodological framework of pragmatic genealogy developed in Chapters 1–3, we can develop the truth of that remark as it applies to Hume.[20] Hume's failure to situate his objects in space and time is no surprise if what he is offering are dynamic models aiming to reveal the point of certain conceptual practices and the way in which that point depends on contingent but highly general facts about us and the kind of environment we live in. What he tries to be true *to* is not the distant past, but the instrumental dependences between certain needs and practices in a certain kind of environment. If his genealogy has any pretensions to realism beyond that, these are psychological rather than historical.

Stripping the state of nature of historical pretensions need not relegate it to the status of mere illustration. A fictional state of nature can still perform several functions. The first is that it serves as a maximally perspicuous and noise-free representation of the problem to which the virtue of justice constitutes a solution, a problem shown to result from the combination of selfishness and partial generosity on the side of the agents with scarcity and instability of goods on the side of the environment. Hence, the state of nature 'deserves our attention, because nothing can more evidently shew the origin' (*T*, 3.2.2.14) of justice in the highly general problem of avoiding conflict over external goods.

A second function of the state of nature is that it allows us to separate in fiction what is inseparable in reality. We can treat 'as compounded and consisting of two parts separate from each other' any 'motion' that we acknowledge to be in fact 'uncompounded and inseparable' (*T*, 3.2.2.14):

> Human nature being composed of two principal parts, which are requisite in all
> its actions, the affections and understanding, it is certain that the blind motions
> of the former, without the direction of the latter, incapacitate men for society;

---

[19] See also Lifschitz (2012). But see Marušić (2017) for a different interpretation of Stewart (1858) on which the aim of conjectural history is not necessarily to speculate about how things actually came about, but to describe the simplest and most natural way that they could have arisen. On this interpretation, conjectural history is less beholden to what actually happened than on the common interpretation, though it remains more beholden to it than Hume's genealogy with its impossible state of nature.

[20] See Neuhouser (2014, 35) for a discussion of why Rousseau's own genealogy of inequality is not best understood as a conjectural history either. Rousseau writes: 'Let us therefore begin by setting aside all the facts...The inquiries that may be pursued regarding this Subject ought not be taken for historical truths, but only for hypothetical and conditional reasonings' (1977, 132).

and it may be allowed us to consider separately the effects that result from the separate operations of these two component parts of the mind.   (*T*, 3.2.2.14)

Hume mentions comparable feats of fictional 'uncompounding' by the 'natural philosophers' whose example he professes to follow. A salient precedent is Galileo Galilei's 1638 analysis of the motion of projectiles in his *Dialogues Concerning Two New Sciences*:

> I now propose to set forth those properties which belong to a body whose motion is compounded of two other motions, namely, one uniform and one naturally accelerated. . . . This is the kind of motion seen in a moving projectile; its origin I conceive to be as follows: Imagine any particle projected along a horizontal plane without friction; then we know . . . that this particle will move along this same plane with a motion which is uniform and perpetual. . . . But if the plane is limited and elevated, then the moving particle . . . will on passing over the edge of the plane acquire, in addition to its previous uniform and perpetual motion, a downward propensity due to its own weight; so that the resulting motion which I call projection is compounded of one which is uniform and horizontal and of another which is vertical and naturally accelerated.   (Galilei 1954, 244)

In this passage, which marks the discovery that projectile motion can be analysed as a combination of two independent linear motions, Galilei's insight is precisely that we can treat 'as compounded and consisting of two parts separate from each other' any 'motion' that we acknowledge to be in fact 'uncompounded and inseparable'. In Hume's *Treatise*, we find the same methodological insight transferred from motions to actions: just as the curved motion of the projectile can be understood as combining an inertial motion in the horizontal direction with a falling motion in the vertical direction, human actions can be understood as combining 'the blind motions' of the affections with the motions of the foreseeing understanding. One complex phenomenon is shown to result from the interplay of several distinct forces, each of which is imagined to play itself out in isolation from the others, the better to understand its contribution to the whole.

On Hume's analysis, the combination of four notionally separable parameters turns out to be crucial to understanding the origin of the virtue of justice, namely the degrees to which (i) agents are given to pursuing their *self-interest*; (ii) agents are given to impartial *benevolence*; (iii) external goods are readily *available*; and (iv) possession of external goods is *stable*. Self-interest and benevolence pertain to the *affections*, while the availability and stability conditions inform the *understanding* in determining how best to satisfy those affections. By allowing us to 'consider separately the effects that result from the separate operations of these two component parts of the mind', the state of nature becomes a heuristic device, a thinking tool that allows us to gain traction on a problem by manipulating our

model of it, exploring what each parameter is doing and what varies with what if we fiddle with the parameters.

This enables the state of nature to perform a third function: to reveal how the pointfulness of the virtue of justice depends on certain contingent facts about the needs and circumstances of creatures like us. In actual fact, humans are strongly given to pursuing their self-interest, their benevolence is limited and partial, and external goods are scarce and unstable.[21] The virtue of justice derives its point from these contingent facts because it solves a problem that only arises when these four parameters take these values. It becomes otiose when the parameters take values outside this range, as Hume shows by turning the knobs of his hypothetical situation to settings with which the problem to which justice answers fails to arise. In a fictional 'golden age' in which either the benevolence of our notional agents or the availability of goods in their environment are ramped up, the challenge at the origin of the virtue of justice disappears:

> [I]f every man had a tender regard for another, or if nature supplied abundantly all our wants and desires, [then] the jealousy of interest, which justice supposes, could no longer have place; nor would there be any occasion for those distinctions and limits of property and possession, which at present are in use among mankind. Encrease to a sufficient degree the benevolence of men, or the bounty of nature, and you render justice useless... (*T*, 3.2.2.16)

Justice has its home in a world of scarcity and limited generosity. Paradisiac circumstances render it quite pointless. As Heraclitus remarked, were it not for quarrels and conflicts, humans would not have known the name of justice.[22] But justice becomes equally pointless at the other extreme of a fictional 'war of all against all' marked by 'untamed selfishness and barbarity' and 'attended with the most extreme necessity' (*EPM*, 3.1.16): 'if such a state of mutual war and violence was ever real, the suspension of all laws of justice, from their absolute inutility, is a necessary and infallible consequence' (*EPM*, 3.1.16).

The upshot of this exercise in counterfactual thinking is twofold. On the one hand, it reveals *limits* to the range of circumstances under which the virtue of justice has a point. In the extreme scenarios of a Hobbesian 'war of all against all' or of a 'golden age', justice would have no intelligible role to play. On the other hand, it also shows that range to be *broad enough* to justify the expectation that any actual human society will cultivate something like the virtue of justice. On Hume's account, any human society that wants to subsist between the twin

---

[21] These natural endowments function as unexplained explainers: 'It is needless to push our researches so far as to ask, why we have humanity or a fellow-feeling with others.... We must stop somewhere in our examination of causes' (*EPM*, 5.2.2n19).
[22] Heraclitus (1981, fragment 23). See Williams (1995e, 137–8).

extremes of paradisiac abundance and total war will find the virtue of justice worth having.

Significantly, Hume thereby reveals the entrenched contrast between the *innate and invariable* on the one hand and the *acquired and variable* on the other to be simplistic. He identifies conceptual practices that are not innate, but that we should nevertheless not expect to vary much, as they are held in place by *structural needs*. Structural needs are second-order needs that derive from the relations between the first-order needs of individuals (such as their need for various types of foods, goods, and tools). Structural needs are counterfactually robust insofar as they are insensitive to the content of first-order needs. Whenever the satisfaction of first-order needs is competitive because individuals share a living space with finite resources, the structural demands in which the virtue of justice is rooted are bound to arise. As long as these demands persist, they impose a constraining functional mould on possible proprieties of use for external goods. Hume sees similar functional constraints at work elsewhere:

> How great soever the variety of municipal laws, it must be confessed, that their chief out-lines pretty regularly concur; because the purposes, to which they tend, are every where exactly similar. In like manner, all houses have a roof and walls, windows and chimneys; though diversified in their shape, figure, and materials. The purposes of the latter, directed to the conveniences of human life, discover not more plainly their origin from reason and reflection, than do those of the former, which point all to a like end.   (*EPM*, 3.2.24)

However diversified the property conventions of different societies may be, Hume's pragmatic genealogy gives us reason to expect that they *point all to a like end*, namely the end of remedying conflicts over external goods.

How does Hume's genealogy deal with the claim—which has often been made in the service of colonizers' expropriation of the colonized—that some cultures lack a concept of property? Would robust anthropological data to this effect suffice to falsify his genealogy? The first thing to note is that his genealogy informs how one approaches such data. It cautions us against inferring from the fact that a culture lacks *our* concept of property that they must therefore lack a concept of property altogether. What the genealogy encourages us to look for is not something that is similar in *content* to our concept of property, but something that is *functionally* analogous to it: something that performs a similar function by helping to solve the problem of conflict over external goods *as that problem expresses itself in that society*. The genealogy helps us discern such an analogue under the guidance of a generic dependence structure rather than a specific concept of property. By enabling us to see *of which general solution* the eighteenth-century Scottish concept of property is a *particular elaboration under specific circumstances*, Hume's model allows us to make rough predictions as to what different

elaborations we should expect under different circumstances, where both the problem and the solution to it may express themselves in different forms. In Enlightenment Britain, for example, land was a scarce and unstable external good in high demand, so Hume's model predicts that conflict prevention would involve landownership. In the vast spaces of Africa or America, by contrast, the chief source of conflict was often not land, which was abundant enough at the time, but other external goods such as cattle. Hume's model predicts that property norms would receive a very different elaboration under such circumstances—so different, indeed, that it might fail to be recognized as a solution to the same problem.

Of course, even though Hume's model gives us strong reasons to expect any human community to have faced the problem of conflict over external goods and to have developed some solution to it, these reasons remain defeasible. The functional dynamics it highlights might in fact have been overridden by other forces, so that a solution might have failed to arise or been rendered unnecessary. But the mere fact that we do not universally find *our* concept of property does not yet put pressure on Hume's genealogy. What would indicate that something has gone wrong would be data to the effect that even in the absence of obvious defeaters, the practical pressures he identifies systematically fail to give rise to the solutions he anticipates. Hume's genealogy not only needs to be internally validated by the inherent plausibility and tractability of each of its steps; it also needs to be externally validated by its ability to predict and explain the actual configurations of human cultures.

There is, finally, a fourth function the state of nature fulfils in Hume's hands. In a way that echoes Rousseau's use of the device,[23] it serves as a sketch of a *psychological path* through which the virtue of justice might have arisen without making great demands on the understanding or foresight of those involved. Among the genealogists discussed in this book, Hume is the one who makes the most elaborate use of the state of nature as a means of identifying the psychological path through which practical inconveniences might come to be remedied. This fourth function thus merits a closer look.

Hume declares that he pursues 'two questions', one concerning the 'reasons' why we might have come to consider justice a virtue, and another concerning the 'manner, in which rules of justice are established by the artifice of men' (*T*, 3.2.2). Hume thus takes pains to identify not only the rationale for the virtue of justice, but also the string of motives explaining how humans might come to develop and value dispositions of justice without fully understanding the point of doing so.

What are Hume's motives for this inquiry into motives? We can distinguish three. The first is the concern to avoid explanatory circularity, which we encountered

---

[23] Neuhouser (2014) brings out the strong affinities of Rousseau's genealogy of inequality with Hume's use of the state of nature as a model of psychological origins. But see Griswold (2018, 53) for an account of Rousseau's genealogy that contrasts it with Hume's.

already in the Circle Argument and which Hume flags again as he introduces the subject of promising.[24] An explanation in which the people in the state of nature *aim* to bring about a state of affairs in which the virtue of justice is widespread *because it is a virtue* would presuppose part of what it is meant to explain, namely an appreciation of the content of the ideas of justice and property.

The second motive is Hume's concern to avoid what might be called *intellectualism*. A genealogical explanation will be intellectualist if it makes unrealistically strong demands on the faculties of foresight and understanding of the originators of a conceptual practice. The tendency to have an exaggerated sense of the complexity and sophistication of typical human performance is a bias combining anthropocentrism with confabulation which Cameron Buckner aptly terms *anthropofabulation* (2013). As Buckner notes, Hume is one of the first to diagnose this anthropofabulation bias. To counter that bias, Hume makes it a rule not to explain human behaviour in terms that are special to it if it can also be accounted for in terms that apply to the rest of the animal kingdom:

> When any hypothesis...is advanc'd to explain a mental operation, which is common to men and beasts, we must apply the same hypothesis to both;...The common defect of those systems, which philosophers have employ'd to account for the actions of the mind, is, that they suppose such a subtility and refinement of thought, as not only exceeds the capacity of mere animals but even of children and the common people in our own species.   (*T*, 1.3.16.3)

Hume's concern to avoid intellectualist explanations fuelled by anthropofabulation is evident in his insistence that his genealogy of promising draws on a deflationary conception of human capacities:

> Nor is that knowledge, which is requisite to make mankind sensible of this interest in the *institution* and *observance* of promises, to be esteem'd superior to the capacity of human nature, however savage and uncultivated. There needs but a very little practice of the world, to make us perceive all these consequences and advantages. The shortest experience of society discovers them to every mortal.   (*T*, 3.2.5.11)

Hume also avoids drawing on public-mindedness as a motive driving the adoption of artificial virtues: 'men, in the ordinary conduct of life, look not so far as the public interest...That is a motive too remote and too sublime to affect the generality of mankind' (*T*, 3.2.1.11).

---

[24] A promise, he writes, '*wou'd not be intelligible, before human conventions had establish'd it*', and '*even if it were intelligible, it wou'd not be attended with any moral obligation*' (*T*, 3.2.5.1).

Thirdly and *a fortiori*, Hume is concerned to avoid the compounded mistake of combining circularity with intellectualism. His genealogical explanation would be guilty of *circular intellectualism* if it represented the target situation as the outcome of a collective deliberation in which people reasoned that they would be better off if they cultivated the virtue of justice and ensured stable possession of external goods by conceiving of them as forms of property. Not only would this make implausibly strong demands on those individuals' powers of foresight; it would also presuppose part of what it is meant to explain, namely *appreciation of the content* of the ideas of justice and property and the attendant reasons for action, including notably a concern for public interest.

To avoid these pitfalls, an explanation of the emergence of practices that are strongly in the public interest must appeal neither to a concern for the public interest nor to the concepts and motivations that the genealogy is meant to explain. Hume's talk of artificial virtues being products of 'contrivance' might be taken to flout this requirement, as it can seem to imply that artificial virtues are the result of conscious design. But the mere fact that the virtue of justice is artificial in Hume's sense does not entail that it is the result of conscious design. Something can be artificial in that sense by being the result of conscious action but *not* of conscious design. Like the invisible-hand explanations first intimated by Bernard Mandeville's *Fable of the Bees* (1988) and developed by Smith (1977, 6.2.9; 2002, 6.1.10), Hume's genealogy presents agents as 'stumbling upon establishments, which are indeed the result of human action, but not the execution of any human design' (Ferguson 1966, 187).[25] Two key differences between pragmatic genealogies and invisible-hand explanations remain, however. First, according to Edna Ullmann-Margalit, invisible-hand explanations identify a process that acts as 'an aggregate mechanism that takes as "input" the diverse and dispersed actions of numerous individuals, and produces as "output" an overall, structured, social pattern' (1997, 190), but as she highlights, there is no assumption that the pattern in question serves a *function*. In contrast to pragmatic genealogies, then, invisible-hand explanations do not necessarily take a function-first approach. Second, invisible-hand explanations concern what already *looks* designed. They account for the appearance of design without appealing to design, explaining the appearance of design as the unintended product of aggregated intentional actions.[26] Hume's genealogical explanation, by contrast, reveals functionality in a phenomenon that does not even *look* designed. Hence, the very suggestion that the virtues

---

[25] Mandeville writes: 'we often ascribe to the Excellency of Man's Genius, and the Depth of his Penetration, what is in Reality owing to length of Time, and the Experience of many Generations' (1988, II, 104). However, as Ullmann-Margalit (1997, 183) points out, Mandeville only floated the idea *that* social order may arise without design; it is only with Hume and Smith that mechanisms explaining *how* such order might arise are proposed.

[26] See Tieffenbach (2013) for a revisionary account of invisible-hand explanation according to which the appearance of design can be intentional, provided individuals see what they do as an aggregation of their individual actions rather than as something they jointly perform.

of justice or fidelity to promises might serve a function is already the first consequential step in Hume's approach. It elicits a radical aspect-shift in the manner in which we view these virtues by inviting us to consider them from a practical point of view.

It is of course true that there are some deliberative and intentionalist elements in Hume's story, and one needs to ask whether these are dispensable aspects of the genealogy's mode of presentation—aspects that, in Hume's phrase, pertain to the manner rather than the matter—or whether they threaten his claim to escaping circularity and intellectualism. I think the deliberative and intentionalist elements in Hume's story stop short of circularity and intellectualism. This can be shown by comparing two deliberative schemas:

Deliberative Schema A

(P1)   We need to solve the problem of conflict over possessions.

(P2)   Adopting rules that stabilize possessions by turning them into property and cultivating respect for property as a virtue would solve that problem.

(C1)   Therefore, let us adopt these rules and cultivate that virtue.

Deliberative Schema A depicts a collective, intellectualist, and explanatorily circular deliberation. Despite its deliberative and intentionalist elements, the deliberation that Hume's genealogy relies on is clearly different. It is better represented on the model of *Deliberative Schema B*:

Deliberative Schema B

(P1)   I have reason to abstain from the possessions of others, provided they abstain from mine.

(P2)   Others similarly have reason to abstain from my possessions, provided I abstain from theirs.

(P3)   I know that they know this, and they know that I know that they know this, and so on, which means that it is common knowledge.

(C1)   Therefore, I shall tentatively abstain from the possessions of others.

This is all that is required to get justice off the ground. With a similarly modest modicum of reflection and deliberation, the mechanisms of sympathy and education then take care of the rest. Neither the need to remedy certain inconveniences nor the fact that justice can in fact remedy them enter as premises into this deliberative process. The same is true of promising: people 'by concert, enter into a scheme of actions, calculated for common benefit, and agree to be true to their word', but the calculation of the common benefit *does not occur in the content of their deliberation*—nothing more is 'requisite to form this concert or convention,

but that every one have a sense of interest in the faithful fulfilling of engagements, and express that sense to other members of the society' (*T*, 3.2.5.11).

Hume's achievement can be captured in terms of a distinction between *sharing the needs* that give point to a practice and *understanding the point* of the practice: he shows how the fact that humans share certain needs that give the virtue of justice its point can be enough for the virtue to emerge, whether or not they understand its point. To say that they *share* the needs that give the practice its point without necessarily *understanding* that point means: they *would* recognize the practice as an improvement and aim to bring it about *if* they understood the practice's relation to their needs. But since understanding a practice's relation to one's needs presupposes possession of the concept in question, it is crucial to the viability of Hume's explanation that mastery of the concept can be achieved antecedently and independently of understanding the point of the concept. The consequence of neglecting the sharing/understanding distinction would be that a new conceptual practice could only comprehensibly emerge if it had emerged already—that is to say, it could not emerge at all. Hume's genealogy avoids this stumbling block by delineating a genealogical path that does not require understanding the point of the practice, and hence does not presuppose appreciation of its conceptual content in advance of its emergence.

Finally, it is owing to the fact that a conceptual practice can have a point for those who engage in it without that point necessarily being transparent to them that allows Hume's genealogy to affect our attitude towards its subject matter. That is why revealing the genealogical inquiry does not leave everything as it was. By Hume's own lights, the genealogical story *reinforces* our 'sense of morals' (*T*, 3.3.6.3). This sense of morals, he writes, 'must certainly acquire new force when, reflecting on itself, it approves of those principles, from whence it is derived, and finds nothing but what is great and good in its rise and origins' (*T*, 3.3.6.3). As David Wiggins notes, one who accepts the

> Humean account of why he, the subject himself, feels or thinks as he does need not find that, if he accepts the explanation, this subverts his finding of moral beauty in this or that act or character. The acceptance of the explanation will not undermine the reasons he would give himself to follow this or that line of action. For the explanation, the conviction, and the reasons will coexist happily enough in the agent's consciousness.   (2006, 36)

In my terminology, Hume's genealogical explanation is *non-reductive*: it does not unmask reasons for action as really being something else; it does not *identify* some high-minded motive with another, baser motive. On the contrary, it explains how and why genuinely new reasons for action would naturally have arisen, and it shows that there are good reasons for us to think as we do in this respect.

This makes Hume's genealogy vindicatory in three respects. First, it offers a kind of *negative* vindication: 'nothing is presented on any side, but what is laudable and good' (*T*, 3.3.6.3). Hume's genealogical reflection does not reveal anything about the origins of the virtue of justice that is radically incompatible with allegiance to it.[27] Second, it *naturalizes* the potentially puzzling or mysterious idea of justice. It does this by showing that the virtue of justice is firmly rooted in human needs and its emergence can be accounted for without resorting to supernatural or metaphysical explanatory material or to circular appeals to what it seeks to explain: a series of basic human motives provide a straightforward psychological path through which societies lacking the virtue of justice would naturally be led to cultivate it.[28] Third, Hume offers a *pragmatic* vindication of the virtue of justice in that it shows how, if human beings were to reflect on their needs and what they entail in the environment they inhabit, they would recognize it to be *instrumentally rational* for them to promulgate the virtue of justice. Hume shows that justice is practically necessary for society, and society is in turn practically necessary, or very nearly so, for human life. This shows why any society must try to cultivate in its members the disposition to be just, and why, beyond a certain critical mass, it will not survive failure in that respect—there are only so many knaves the institution of justice can take.

Granted that Hume's genealogy provides a *vindication* of the virtue of justice, it is worth asking *for whom* it functions as a vindication. After all, showing that there should not be too many knaves if society is not to disintegrate does not provide the knave with a reason to cease to be one. What we have here is merely an *external* vindication of justice—one that gets no internal hold on the knave's position.[29] The very fact that he is seen *as* a knave shows that he can exist without making the institution of justice collapse. The thought that if everybody were like him, he could not exist will only get a grip, and turn the external justification into an internal one, to the extent that he otherwise shares in the ethical life. The imagined universalization is an essentially moral thought, and the knave will only be receptive to it insofar as he *already is* what we want him to become. In this respect, Hume's justification of justice can do no better than other attempts to justify the ethical: it must preach, as it were, to the choir.[30] But this does not mean that it is useless. Rather, as Williams said of attempts to justify the ethical life, the 'aim is not to control the enemies of the community or its shirkers but, by giving reason to people already disposed to hear it, to help in continually creating a

---

[27] Sauer (2018) argues that most vindicatory moral genealogies operate along these lines—that a moral belief should be seen as vindicated if it does *not* get debunked under genealogical reflection, because this gives it an edge over ideas or beliefs that are vulnerable to genealogical debunking.

[28] Within Hume's empiricist programme, a further desideratum for a naturalistic explanation is that it should be able to show how justice 'is not founded on our ideas, but on our impressions' (*T*, 3.2.2.20).

[29] Williams (1973b, 252–3) draws a similar contrast between external and internal justifications in discussing the justification of altruism.

[30] See Blackburn (2010, 294) and Williams (1973b; 2011, 26).

community held together by that same disposition' (2011, 31). Similarly, Hume's justification of justice is not meant to serve as an instrument of conversion. But it can promote self-understanding, and thereby strengthen the confidence of those who are already somewhat disposed to be just.

More could be said about how Hume's pragmatic genealogical method serves his naturalistic ambitions. Yet given how well explored Hume's naturalism is, the connection between pragmatic genealogy and naturalism can shed more light on a genealogist harbouring similar naturalistic ambitions, but whose resemblance to Hume has until recently been largely overlooked:[31] Friedrich Nietzsche. Bringing this unlikely candidate into the fold of pragmatic genealogy and highlighting the method's continuity with the thought of this preeminent genealogist is the task of Chapter 5.

---

[31] See Stroud (2006), Russell (2008), and Kail (2007b, 2009, 2016) for recent discussions of Hume's naturalism. For comparisons between Hume and Nietzsche, see Hoy (1994), Beam (1996, 2001), Broisson (2006), and Kail (2009, 2016).

5

A Genetic History of Thought

Friedrich Nietzsche

'The thinker', Nietzsche writes, 'regards everything as having evolved...he asks: whence does it come? what is its purpose?' (*WS*, §43). This programmatic statement encapsulates the triad of naturalism, genealogy, and pragmatism in Nietzsche's thought that forms the focus of this chapter. From the naturalistic standpoint from which everything is seen as having evolved, understanding is sought by inquiring into the origins and purposes of things. After Darwin, this became the standard approach to the traits of living things. Nietzsche's contribution lies in systematically extending it to *concepts*. He is the first genealogist to explicitly envisage a systematic deployment of the genealogical method across the entire range of thought, calling on scientists and scholars to draw on the 'physiology and history of the evolution of organisms and concepts' in order to celebrate their 'highest triumph in a *genetic history of thought*' (*HA*, §16).

There is reason to think that this project of offering a genetic history of thought can also be fruitfully viewed through the prism of pragmatic genealogy. From the 1870s onwards, Nietzsche makes sense of metaphysics-inviting concepts that appear unconditioned by history and mundane practical concerns as the historical outgrowths of human needs: 'Unsere *Begriffe* sind von unserer *Bedürftigkeit* inspirirt' (*eKGWB*, 1885, 2[77])—our concepts are inspired by our need. If we look at the '*true genesis of concepts*', we shall find that they stem from the 'practical sphere' (*eKGWB*, 1887, 8[2]).

Reading the early Nietzsche in light of the pragmatic genealogical framework throws several neglected aspects of his work into relief. First, we find that he already sketches genealogical explanations of truthfulness and justice during his Basel period (1869–79). Second, these genealogies are not primarily documentary, subversive, and revelatory of contingency, as Foucault's reading of Nietzsche would lead one to expect. Rather, they are primarily imaginary, vindicatory, and revelatory of practical necessity.[1] Third, this brings the early Nietzsche much closer to the 'English genealogists' than his tirade against them at the beginning

---

[1] I say 'primarily' because the contrast—drawn in one form or another in Paden (2003, 566), Jenkins (2006, 164), Hoy (2009, 225), Koopman (2009), and Millgram (2009, 163n23)—between genealogies that are documentary, subversive, and revelatory of contingency and genealogies that are imaginary, vindicatory, and revelatory of practical necessity is not best treated as an exclusive one. Nietzsche's oeuvre alone combines all these aspects in varying proportions over his lifetime.

of *On the Genealogy of Morality* would encourage one to think.² Genealogies in this 'English' style, the mature Nietzsche tells us, seek to naturalize values through functionalist hypotheses about why they might have originated: they start out from a hypothesis about the original function of a certain way of valuing, and then suggest that while the valuation solidifies through habit, its function is forgotten, so that it is unjustifiably extended beyond the boundaries of its original functionality (*GM*, I, §2). We will see that it is along just these lines that Nietzsche himself proceeds in his early genealogies of truthfulness and justice. Pragmatic genealogy thus turns out to have as much of a claim to being of Nietzschean descent as the better-known Foucauldian variety.

An advantage of viewing Nietzschean genealogy through the prism of pragmatic genealogy is that it accounts for his lack of historical references. Most interpreters tend to read Nietzschean genealogy as being in effect some form of historiography; yet the problem for these interpretations is that they have a hard time accounting for what can only appear, in a trained philologist, as a poor effort at writing history. This problem, real enough in the *Genealogy*, is even more acute in his earlier genealogies: the only historical reference in his genealogy of justice, for example, is to a situation in which questions of justice precisely *failed* to arise. If, by contrast, we read Nietzsche as trying to identify the point of justice using a fictional model comparable to what other pragmatic genealogists call a 'state of nature', his failure to engage with history becomes comprehensible.

I begin by showing how Nietzsche's genealogical method can be seen as a corrective to what he saw as philosophers' deplorable tendencies towards dehistoricization and denaturalization before focusing on Nietzsche's application of the genealogical method to concepts. I argue that he is concerned to present seemingly unconditioned concepts as being in fact conditioned by the contingencies of history and the demands of human needs. I then offer a detailed examination of two of Nietzsche's early genealogies, his genealogy of justice and his genealogy of truthfulness, and highlight the distinctive contributions Nietzsche makes to the tradition of pragmatic genealogy.

---

² In the First Treatise of the *GM*, Nietzsche repeatedly refers to the 'English psychologists', and in Höffding's *Psychologie in Umrissen auf Grundlage der Erfahrung* (1887, 56, 61, 195), which Nietzsche drew on, 'English psychology' refers to a school of psychology that privileges mechanistic explanations in terms of association and habit (Sommer 2019, 88–90). In the *GM*'s Preface, moreover, Nietzsche presents his book as an answer to Paul Rée and other 'English genealogists of morality', which may be a reference to the protagonists of W. E. H. Lecky's *History of European Morals* (1869), of which Nietzsche owned a translation he annotated extensively. In the chapter entitled 'A Natural History of Morals', Lecky outlines the views of Bain, Bentham, Hartley, Hobbes, Hume, Hutcheson, Locke, Mandeville, Mill, Paley, Shaftesbury, Smith, Spencer, and others—see Thatcher (1989, 588) and Brobjer (2008a, b). The term 'English' in Nietzsche's usage seems to refer primarily to an intellectual style, since it covers both Hume (who was Scottish) and Rée (who was German). Juxtaposed with Rée's, Nietzsche's book comes off as an attempt to improve on it point by point, which suggests that Rée was foremost on Nietzsche's mind (Hufendiek 2019; Janaway 2007, 78).

5.1 Philosophers' Dehistoricizing and Denaturalizing Tendencies

Philosophers, Nietzsche finds, suffer from two idiosyncrasies that are detrimental to philosophy: they tend to *dehistoricize* and *denaturalize* what they respect. Each idiosyncrasy can express itself in three different ways. The tendency to dehistoricize can take the form of (a) insensitivity to historicity, (b) the perception of historicity as a defect in what one respects, or (c) the desire to veil historicity by thinking of what one respects in ahistorical terms. The tendency to denaturalize, meanwhile, characteristically gives rise to such convictions as (d) that the lesser cannot give rise to the greater, (e) that the greater must be made sense of in terms that demand more material, such as theological or metaphysical stories of creation or revelation, or (f) that the greater must arise out of itself, as a *causa sui*.

   Writing downstream of the advent of historicism and Darwinian naturalism, Nietzsche seeks to correct these twin tendencies towards dehistoricization and denaturalization with his genealogical method. Historicization and naturalization are sometimes conceived of as mutually exclusive movements of thought; but Nietzsche's genealogical method unproblematically combines the two: it is historicizing in its recognition of the need for historical understanding, in its acknowledgement of the omnipresence of change, and in its re-historicization of what has been dehistoricized; and it is naturalizing in that it explains how the lesser can give rise to the greater, thereby making appeals to revelation, creation or self-causation redundant, and rendering intelligible how highly abstract, highly complex, or highly valued cultural formations have their roots in natural drives and needs. Let us take a closer look at what Nietzsche has to say about each tendency.

5.1.1 The Tendency to Dehistoricize

The dehistoricizing tendencies of philosophers are denounced by Nietzsche already in 1878: 'Lack of a historical sense is the hereditary defect of philosophers …So what is needed from now on is *historical philosophizing*, and with it the virtue of modesty' (*HA*, §2). There is considerable irony in this formulation, since historical sense is precisely what is required to recognize something as a *hereditary* defect ('*Erbfehler*'). To the extent that it must itself be conceived of in historical terms, lack of historical sense will therefore tend to be self-concealing, leaving those most afflicted by it least qualified to perceive their own shortcomings. Consequently, lack of historical sense has incrusted itself as an 'idiosyncrasy' of philosophers:

> You want to know what the philosophers' idiosyncrasies are?…Their lack of historical sense for one thing, their hatred of the very idea of becoming, their

Egyptianism. They think they are showing *respect* for something when they dehistoricize it, *sub specie aeterni*,—when they turn it into a mummy. For thousands of years, philosophers have been using only mummified concepts; nothing real makes it through their hands alive.   (*TI*, Reason, §1)

As this passage indicates, the tendency to dehistoricize has three aspects. The first is the lack of historical sense, which blinds one to historicity and leads one to dehistoricize out of ignorance. It is not just the 'English psychologists' (*GM*, I, §1) who lack historical sense on his view, but also the German philosophers of his time. 'Philosophers', Nietzsche observes, 'all think *essentially* ahistorically' (*GM*, I, §2). Chiding Schopenhauer for thinking that the history of morality could be traced to a single origin in pity, he writes that this was possible only 'for a thinker who was robbed of all historical instinct and who had in the strangest way even escaped that strong schooling in history that the Germans have been through from Herder to Hegel' (*eKGWB*, 1885, 2[188]). Nietzsche even blames Schopenhauer's rants against Hegel for causing 'the whole of the last generation of Germans to break off its ties to German culture, a culture that, all things considered, represented a supreme and divinatory refinement of the *historical sense*' (*BGE*, §204).

Apart from expressing itself in a lack of historical sense, however, philosophers' tendency to dehistoricize is also evinced in their 'hatred of the very idea of becoming', which is an evaluative attitude towards historicity: philosophers see 'death, change, and age, as well as procreation and growth' (*TI*, Reason, §1) as objections, flaws, imperfections to be avoided. If being subject to change and becoming are flaws, it follows that truly respectable concepts must be exempt from such flaws: 'all the highest concepts, Being, the Unconditioned, the Good, the True, the Perfect—none of these could have become' (*TI*, Reason, §4). The best and highest should not be subject to change at all.

The third aspect of the tendency to dehistoricize is what Nietzsche dubs *Egyptianism*: the tendency to express one's respect for concepts by *mummifying* them. Concept mummification consists in the attempt to deny the historicity of concepts by treating them as eternal and unchanging—one develops ahistorical *conceptions* of what are in fact historical concepts. Mummification involves a form of self-deception: one deceives oneself into thinking of what one respects as something eternal, thus suppressing or silencing its historicity.

For Nietzsche, these three aspects of the dehistoricizing tendency of philosophers have fuelled a denial of the historical world accessible to the senses and a striving for an ahistorical, unconditioned and eternal world beyond it:

What is, does not *become*; what becomes, *is* not … So they all believe, desperately even, in being. But since they cannot get hold of it, they look for reasons why it is kept from them. 'There must be some deception here, some illusory level of appearances preventing us from perceiving things that have being: where is the

deceiver?'—'We've got it!' they shout in ecstasy, 'it is sensibility! These senses *that are so immoral anyway*, now they are deceiving us about the *true* world. Moral: get rid of sense-deception, becoming, history, lies—history is nothing but a belief in the senses, a belief in lies'.   (*TI*, Reason, §1)

From Plato's eternal Forms through the Christian God to the noumenal world and the thing in itself—all of them, Nietzsche thinks, have served as timeless realities, homes to 'higher' things relative to which the 'lower' historical world of the senses along with its perceived imperfections could be demoted to the status of a mere 'illusion' or 'lie'. Nietzsche in effect agrees that the unconditioned ideals envisaged by the philosophical tradition are incompatible with the characteristics of the historical world. But while his predecessors took this as an objection to the historical world, he takes it as an objection to their ideals: 'In a world of becoming in which everything is conditional, the assumption of the unconditional . . . can only be error' (*eKGWB*, 1885, 35[51]).

5.1.2 The Tendency to Denaturalize

If one idiosyncrasy of philosophers is their conviction that the highly complex and highly regarded should not change or grow at all, the other, Nietzsche tells us, is that it should certainly '*not* grow out of the lowest' (*TI*, Reason, §4). But *can* something emerge from its opposite, at least in the weak sense in which X emerges from non-X? This is a question at the heart of Nietzsche's concerns, and he takes it to have been at the heart of philosophy itself for the past 2,000 years. Accordingly, he opens not only *Human, All Too Human* but also *Beyond Good and Evil* with precisely this question. The issue is fundamentally that of the merits of *naturalism*: can we find a place for such abstract and ethereal phenomena as rationality, life, logic, altruism, truthfulness, and justice in the natural world by explaining how they could have emerged out of the rough-and-tumble of a reality originally devoid of these things?

   Philosophers, Nietzsche finds, have tended to answer this question in the negative.[3] Instead of trying to explain how a highly valued X could emerge from a lesser non-X, 'metaphysical philosophy has hitherto surmounted this difficulty by denying that the one originates in the other' (*HA*, §1). This attitude, which Nietzsche dubs 'the metaphysicians' basic faith, *the faith in the opposition of values*' (*BGE*, §2), leaves the origins of higher things shrouded in mystery, satisfying 'the demand in the souls of all religious people and metaphysicians (in those of the artists, too, when they are also thinkers)' that the unexplained be

---

[3] See Nietzsche (*HA*, §1; *TI*, Reason, §5).

'altogether inexplicable, the inexplicable altogether unnatural, supernatural, miraculous' (*HA*, §136). Attempts to trace back the 'supposedly miraculous' to the 'complex, the multiply caused' (*HA*, §136) are resisted on the principle that *the lesser cannot give rise to the greater.*

Allegiance to this principle is the first and central aspect of the tendency to denaturalize. Philosophers consider it an objection for the greater to come from the lesser, because 'it casts doubt on its value' (*TI*, Reason, §2)—or so they think.[4] Again, Nietzsche thinks that 'this is just their way of showing respect' (*TI*, Reason, §4).

If the highly abstract, complex, or valued cannot be accounted for in terms of lesser material because a bottom-up direction of explanation is excluded on the grounds of being disrespectful, this leaves only two other directions of explanation: top-down explanations in terms of something even more highly abstract, complex, or valued, or circular explanations in terms of the explanandum itself. It is to these remaining strategies that the second and third aspect of the tendency to denaturalize correspond.

The second aspect consists in explaining the origins of the greater in terms that demand even more material than it itself provides. This is what drives theological or metaphysical revelation stories tracing things to Divine Commands, God-given Tables of Values, or Platonic Forms. It long seemed natural to assume that exalted things could only have derived from even more exalted things—this, Nietzsche argues, is how people ultimately got the idea of the highest and most real of things, the *ens realissimum*, 'their stupendous concept of "God"' (*TI*, Reason, §4).

Since the idea of the *ens realissimum* leaves top-down explanations no more material to work with, this leads to the strategy of explaining the greater in terms of itself, which is the third aspect of denaturalization. It consists in the circular strategy of *causa sui* accounts. Such denaturalization occurs not only with the idea of God, Nietzsche thinks, but whenever one takes what in fact emerged only at the end of a long process of development and projects it back to the beginning, seeing it as unfolding through time but in principle always already present and not brought into being by anything other than itself (*TI*, Reason, §4). All three aspects of the tendency to denaturalize stem from confusing 'what comes first with what comes last' (*TI*, Reason, §4): what comes first in the *order of value* is put first in the *order of causation*, when in fact what comes first in the order of value tends to come last in the order of causation.

In sum, Nietzsche takes there to be two ways of expressing respect which are typical of philosophers but detrimental to philosophy: one is to think that exalted things should not have *become* at all. The other is to think that *if* they have become, they should *not* have become out of lesser things.

---

[4] What Nietzsche himself thinks about how 'low' origins should affect our perception of value is the topic of Queloz and Cueni (2019).

5.2 Concepts Conditioned by History and Functionality

Acting as a corrective to the twin tendencies to dehistoricize and denaturalize, Nietzsche's genealogies reveal concepts we think of as *unconditioned* by history and functionality to be in fact thus conditioned: they have emerged and developed, and they have done so in response to certain needs and interests.

In presenting concepts as conditioned by history and functionality, Nietzsche took himself to do for concepts what Darwin had done for organisms. 'If there is something new in Nietzsche's use of genealogy', Maudemarie Clark writes, 'it is the suggestion that concepts are formed in the same way as other living things' (2015a, 31). There is even a fragment in the *Nachlass* in which Nietzsche explicitly models concepts on living things, comparing them in particular to cells.[5] Drawing on the knowledge of cytology he acquired through his reading of Wilhelm Roux's *Der Kampf der Theile im Organismus* (1881), Nietzsche comes to think of concepts as having a firm core and a more fluid body:

> Concepts are living things, and hence they may grow at one time and shrink at another: many concepts have died a miserable death. They are comparable to cells, with a cell core encased in a body which is not solid…
>
> (*eKGWB*, 1885, 40[51])

For living things, a useful trait's emergence will often be accidental (a random mutation); but its stability and spread through a population will often be non-accidental: it will endure and spread because it makes a useful difference.

Nietzsche accounts for the spread of concepts by the same logic: there is variation of cultural formations in the course of history, and the usefulness of a given formation explains its retention. Although Nietzsche repeatedly describes concepts as 'falsifying' reality because they originate in the obfuscation of differences—a 'concept', he writes, 'is an invention which nothing corresponds to *wholly* but many things *slightly*'—he also takes these abstractions from the individuality and inconstancy of the flux of becoming to be life-serving, because 'with this invented and rigid world of concepts and numbers, man gains a means

---

[5] Nietzsche left behind an extensive corpus of unpublished material at the time of his mental collapse in 1889. Against his wish to have these notes destroyed, his sister published a tendentious and heavily 'edited' selection of them under the title of *The Will to Power*, presenting them as Nietzsche's long-planned *magnum opus*. While Nietzsche had indeed worked on a book of that title, there is reason to think that he abandoned the project before his descent into madness—though see Huang (2019) for a nuanced discussion. Some of the ideas he retained found their way into the *Anti-Christ* (Young 2010, 534–49). The material in the notebooks therefore does not contain the unpublished *summa* of his thought; nor can it claim to be as authoritative as the material he chose to publish. But it would be an overreaction to deny that if used judiciously, this rich corpus of sometimes more cogent because less compressed and stylized notes can shed light on Nietzsche's published writings in a variety of ways. See Reginster (2006, 16–20) for a compelling defence of this view.

of seizing by signs, as it were, huge quantities of facts and inscribing them in his memory' (*eKGWB*, 1885, 34[131]). Depending on what our needs are and how the world is at any given time, some concepts will be more serviceable than others. Over time, the concepts that pay rent will persist and spread, while those that do not will go out of business: 'the *most useful* concepts have remained; however wrong their origin may have been' (*eKGWB*, 1885, 34[63]).

On the basis of this pragmatist picture of concepts, Nietzsche seeks genealogical explanations that explain how the greater could emerge out of the lesser and render both *causa sui* accounts and metaphysical or theological top-down explanations superfluous, thus nipping in the bud the temptation to denaturalize. As Bertrand Russell said of witchcraft: it was never refuted, but it ceased to be interesting.[6] Similarly, Nietzsche does not refute explanations in terms of self-causation or in terms of metaphysical or theological revelation, but he undermines the explanatory need for them by offering more parsimonious counterproposals (*D*, §95).

Nietzsche's ambition to add as little explanatory material as possible—either on the objective side of what the world contains or on the subjective side of agents' capacities of foresight and comprehension—forms one of his enduring methodological tenets: the 'law of parsimony' (*eKGWB*, 1872, 23[30]). 'Method', he maintains in *Beyond Good and Evil*, 'must essentially be the economy of principles' (*BGE*, §13). He already meditates at length on the law of parsimony in his 1872 notebooks: 'the hypothesis which deploys the smallest number of presuppositions and means to explain the world takes precedence over all rivals' (*eKGWB*, 1872, 23[30]). Explanations in terms of 'simpler and better understood forces, especially of the mechanical sort' (*eKGWB*, 1872, 23[30]), should be given precedence over explanations in terms of more complex or less understood forces.

What shines through here is Nietzsche's naturalism: the genealogies of his Basel years answer to a nineteenth-century naturalist's concern that, without metaphysical postulates, some topics of discourse are hard to make sense of in naturalistic terms. Genealogy remedies this by offering a diachronic translation back into nature of what appeared to be beyond the naturalist's grasp. As Williams observed in his classic paper 'Nietzsche's Minimalist Moral Psychology' (2006c), Nietzsche's is an attractive brand of naturalism. Williams himself subscribes to it (1995d, 204), as will become evident in Chapter 7; but I believe it offers a good articulation of the naturalism animating pragmatic genealogies more generally. What is attractive about this brand of naturalism is that it avoids the difficulty that many articulations of naturalism have, of being either too restrictive or too accommodating. It manages to avoid this difficulty because it does not work from the '-ism' down, but rather from the concrete phenomenon up: instead of

---

[6] The line is attributed to Russell by Williams (2014b, 378).

starting from a sharp delimitation of *the realm of nature* into which respectable phenomena must be shown to fit, or from some litmus-test-like *doctrine* of naturalism that yields contradictions with other beliefs and thus marks them out as unpalatable for the naturalist, Nietzsche's naturalism involves taking up a particular *stance* towards phenomena—in Van Fraassen's (2002) sense of approaching phenomena in a way that is expressive of one's evaluative attitudes and epistemic preferences and strategies as well as of one's beliefs.

We can characterize Nietzsche's naturalistic stance—drawing on a variety of dispersed but suggestive remarks of Williams's (1995d, 2000, 2002, 2006c)—as consisting notably of two epistemic-cum-evaluative attitudes, which we can label *minimalism* and *realism*:

Minimalism:

Explain X as far as possible in terms used anyway elsewhere;

Realism:

Appeal first to terms that an experienced, scientifically informed, perceptive, truthful, and unoptimistic interpreter would use.

Nietzsche's minimalism does not carry with it the demand that the material in terms of which explanations are given be the same in every case, as it would be if we tried to explain everything in terms of physics. Rather, what one takes as given in one's genealogical explanation will *change from one case to the next*: moral psychology is explained in terms of non-moral psychology, conscious processes in terms of non-conscious processes, psychology in terms of physiology. Williams later called this the *creeping barrage* conception of naturalism, in reference to the WWI battle tactic whereby the barrage of artillery fire would creep along with the advancing infantry, veiling it behind a thick curtain of smoke and staying just ahead of it in order to avoid a potentially lethal time-lag between the covering fire and the infantry attack (2000, 150; 2002, 23). On the creeping barrage conception of naturalism, the class of *explanantia* moves along with the class of *explananda*, so that one class of things is explained in terms of the next lower class rather than in terms of a fixed base-level class such as 'entities recognized by fundamental physics'; moreover, what acts as *explanans* in one genealogy can act as *explanandum* in another.

By itself, however, minimalism does not offer enough guidance to steer clear of vacuity and reductionism. If we seek to explain an aspect of human behaviour in terms we use anyway for the rest of human behaviour, what terms will be available will in turn depend on one's interpretation of human behaviour, and that interpretation might be, in Nietzsche's marvellous phrase, an expression of 'noble childishness' (*BGE*, P)—or it might be scientifically reductive, mystical, ascetic, or Panglossian. This is why even the roughest characterization of a naturalistic

stance requires a second attitude. We need 'some guiding sense of what materials we should use in giving our economical explanations' if we are to decide whether minimalism is meant to be 'blandly accommodating, or fiercely reductive, or something in between' (Williams 2006c, 306–7). It will be blandly accommodating if it permits the understanding of phenomena such as morality in terms that they themselves invite but which apply only to themselves: among morality's self-conceptions, for instance, is the idea that humans are naturally able to intuit the structure of moral reality. This provides an economical explanation of morality, but it renders puzzling its relation to the rest of nature. If this is what minimalism means, it excludes too little. If, however, we require the terms in which the explanation is given to apply equally to every other domain of nature, minimalism becomes fiercely reductive: we are led towards describing everything in the terms of physics. If this is what minimalism means, it excludes too much.

Hence the second epistemic-cum-evaluative attitude, which, following Williams (2006c, 302), I have called *realism*: the idea that in explaining a particular aspect of human behaviour, one should appeal first to what an experienced, scientifically informed, perceptive, truthful, and unoptimistic interpreter might make of human behaviour. This is of course not a formula. It 'invites one into a perspective, and to some extent a tradition' (2006c, 302) marked by such authors as Thucydides, Diderot, and Stendhal. This is the streak in Nietzsche's writings that attracts the application of Ricœur's phrase, 'the hermeneutics of suspicion'. It is a component of the naturalistic stance that is motivated not so much by a desire for parsimony as by the sense that there is reason for suspicion—'that stories human beings tell themselves about the ethical tend to be optimistic, self-serving, superstitious, vengeful, or otherwise not what they seem to be' (Williams 1995d, 204). To renounce the metaphysicians' faith in the opposition of values is to recognize that the higher and the lower intermesh, and that the latter will help us understand the former, even if it unflatteringly undermines our picture of ourselves as purely selfless or radically unlike other animals.

In making sense of concepts that seem to transcend quotidian practical concerns by revealing those concepts to have developed in answer to such concerns, Nietzsche's genealogical method thus expresses a naturalistic stance. The concepts of a community are explainable in terms of its needs, and *differences* in concepts are explainable in terms of differences in needs:

> Wherever we encounter a morality, we find an evaluation and ranking of human drives and actions. These evaluations and rankings are always the expression of the needs of a community and herd. . . . Since the conditions for preserving one community have been very different from those of another community, there have been very different moralities; and in view of essential changes in herds and communities, states and societies that are yet to come, one can prophesy that there will yet be very divergent moralities.   (*GS*, §116)

To put it cursorily: Nietzsche seeks to naturalize seemingly unconditioned concepts by presenting them as functional entities conditioned by our needs, and these needs in turn as conditioned by historical circumstances. We can see how this would have led Nietzsche to something like the pragmatic genealogical method as a way of reverse-engineering the practical origins of our concepts in our needs.

5.3 Nietzsche's Vindicatory English Genealogies

The clearest examples of pragmatic genealogies in Nietzsche's work are his early genealogical explanations of justice and truthfulness from the 1870s. While the latter genealogy is scattered over many passages, the former is succinctly encased in a section of *Human, All Too Human*, so let us begin with the genealogy of justice.

5.3.1 Nietzsche's Early Genealogy of Justice

In a chapter entitled 'Of the History of the Moral Sensations', we find what both the declared theme of the chapter and the section heading suggest is a genealogy of justice:

> *Origin of justice.*—Justice (fairness [*Billigkeit*]) originates between parties of approximately *equal power*, as Thucydides correctly grasped (in the terrible colloquy between the Athenian and Melian ambassadors): where there is no clearly recognizable superiority of force and a contest would result in mutual injury producing no decisive outcome the idea arises of coming to an understanding and negotiating over one another's demands: the characteristic of *exchange* is the original characteristic of justice [*Der Charakter des Tauschs ist der anfängliche Charakter der Gerechtigkeit*]. Each satisfies the other, inasmuch as each acquires what he values more than the other does. One gives to the other what he wants to have, to be henceforth his own, and in return receives what one oneself desires. Justice is thus requital and exchange under the presupposition of an approximately equal power position: revenge therefore belongs originally within the domain of justice, it is an exchange. Gratitude likewise.—Justice goes back naturally to the viewpoint of reasonable self-preservation, thus to the egoism of the reflection: 'to what end should I injure myself uselessly and perhaps even then not achieve my goal?'—so much for the *origin* of justice. Since, in accordance with their intellectual habit, humans have *forgotten* the original purpose of so-called just and fair actions, and especially because children have for millennia been trained to admire and imitate such actions, it has gradually

come to appear that a just action is an unegoistic one: but it is on this appearance that the high value accorded it depends; and this high value is, moreover, continually increasing, as all valuations do: for something highly valued is striven for, imitated, multiplied through sacrifice, and grows as the worth of the toil and zeal expended by each individual is added to the worth of the valued thing.—How little moral would the world appear without forgetfulness! A poet could say that God has placed forgetfulness as a doorkeeper on the threshold of the temple of human dignity.   (*HA*, §92)

As Nietzsche's dashes indicate, this section is divided into four parts:

(1) *The Emergence and Original Function of Justice*: The first step in Nietzsche's approach is so obvious as to be easily missed, but it already does some of the work—it is to *ask* after the origin of justice, and thereby to historicize a notion that, most evidently in the natural law tradition, presents itself as ahistorical. Against this tradition, Nietzsche maintains that 'there is no such thing as eternal justice' (*HA*, §53). Justice has origins, and Nietzsche's aim is to explain these origins so 'that it can be perfectly understood without the postulation of *metaphysical interference*' (*HA*, §10). Justice is, as the book's programmatic title has it, *human, all too human*. 'It is', as he later puts it, we 'who really and continually *make* something that is not yet there' (*GS*, §301). Accordingly, Nietzsche's investigation into the origin of justice centres on human agents and the practical imperatives that drove them to develop the concept of justice.

The concept of justice, Nietzsche tells us, has its origins in the concurrence of the following conditions:

(C1)   *Equilibrium of Power*: two parties *A* and *B* under circumstances *C* are so well matched that outright pugnacity would result in a drawn-out feud and mutual harm, leaving the victor so badly mauled as to render the spoils of victory useless.

(C2)   *Conflict of Interests*: two parties *A* and *B* under circumstances *C* have interests such that neither party can freely pursue its interests without frustrating those of the other.

When these two conditions are met, the most reasonable resolution of the situation is a *settlement* through exchange. This requires the identification of the specific exchange in which 'each satisfies the other, inasmuch as each acquires what he values more than the other does'. Nietzsche's suggestion is that the concept of justice is rooted in this need to negotiate a settlement between equally powerful parties: justice is originally trade-like, taking the form of an exchange that is satisfactory to both parties. On this account, the original function of justice is to resolve stand-offs between parties of equal power in a manner advantageous

to both. Lack of satisfaction might then generate a demand for *restorative justice* towards the injured party and *retributive justice* towards the injuring party. Hence Nietzsche's suggestion that 'revenge therefore belongs originally within the domain of justice, it is an exchange', and 'gratitude likewise': gratitude arises when one is given *more* than would be just, the desire for revenge when one is given *less*. This explains why Nietzsche goes on to write that justice is not only 'exchange' but also 'requital'.

If (C2) is not met because *A* and *B* have non-conflicting interests, there can be no question of identifying an exchange that is acceptable to both parties, since they can satisfy their interests without it. Similarly, if (C1) is not met, the stronger party will take what it wants and the weaker will have to put up with it. This is the force of Nietzsche's reference to Thucydides: Thucydides describes how Athens sought to conquer the island of Melos (2013, V 84–116). The Athenians sent emissaries to the rulers of Melos and offered them an ultimatum: surrender and pay tribute to Athens, or be destroyed. The Athenians refused to argue over the justice of the situation, because 'in the human sphere judgements about justice are relevant only between those with an equal power to enforce it . . . the possibilities are defined by what the strong do and the weak accept' (2013, 5.89). Justice cannot have originated in interactions between parties of unequal power, because, as Richard Crawley rendered the same passage in his 1874 translation, 'the strong do what they can and the weak suffer what they must' (1998, V 89).

In the second volume of *HA*, Nietzsche goes on to sketch a reason to think that (C1) will usually be met. He calls it the *principle of equilibrium*: fear of dangerous neighbours will systematically drive individuals to form a community and to 'bring its power of defence and attack up to precisely the point at which the power possessed by its dangerous neighbour stands and then to give him to understand that the scales are now evenly balanced'; the community, on this picture, 'is originally the organization of the weak for the production of an *equilibrium* with powers that threaten it with danger', and this equilibrium 'is the basis of justice' (*WS*, §22).

This Thucydidean insight presumably also underlies Nietzsche's remark in the *Genealogy* that equilibrium is the presupposition of all contracts (*GM*, I, §4). Justice originates 'as a settlement between approximately equal powers' (*GM*, I, §4), as a means of preserving equilibrium instead of engaging in a costly fight. But *when* did it emerge? Nietzsche's only historical reference is to a case where issues of justice failed to arise. Clearly, Nietzsche is not primarily interested in the specifics of the historical situation in which justice first arose. He is concerned to identify the practical needs at the root of justice, and to this end, he operates with an imaginary, generic situation that functions like the 'state-of-nature' fiction in Hume's genealogy of justice.

(2) *The Original Motive to Justice*: What makes it reasonable for both parties to negotiate is the need for self-preservation. As Nietzsche implies by speaking of

'reasonable' self-preservation (*einsichtige Selbsterhaltung*), both parties must be reasonable enough to recognize that an exchange is their best bet. Suicidal agents would presumably forfeit the opportunity to resolve conflicts of interests through exchange, and so would non-suicidal agents too unreasonable to assess the situation correctly. Tracing justice to self-preservation makes the account non-circular: the motive from which the two parties in the genealogy originally settle for what come to be regarded as 'just' terms of exchange is not that they value justice, but that they desire their own self-preservation. This does not mean that Nietzsche equates this original motive with the motive that *now* motivates just acts, as he makes clear in a revealing passage he dictated to Heinrich Köselitz:

> General happiness or general love of one's neighbour are *results*, which *may* (or may not!) be attainable through the continual growth of morality. Not to let go of any human achievements and to hold on to the current heights of humanity, this may be a *consequence* of general morality (a side-effect); but what drives people to moral acts, *drives* them now, is not those results, and much less these consequences; neither is it what originally generated recognition of moral predicates. The origin of morality cannot lie in the domain of the moral. One must therefore distinguish: *first*, the results of morality, *second*, the consequences of morality, *third*, the motives of moral acts, *fourth*, the motives of the genesis of moral concepts.   (*eKGWB*, 1880, 3[122])

Nietzsche's insistence that the origin of morality cannot lie in the domain of the moral echoes Hume's non-circularity requirement. Moreover, by presenting justice as emerging out of a drive to self-preservation he takes to be effective in the rest of the animal kingdom as well, Nietzsche heeds the demands of parsimony in a way that Hume would have approved of: 'The beginnings of justice', Nietzsche writes, 'are *animal*: a consequence of that drive which teaches us to seek food and elude enemies' (*D*, §26).

(3) *Emergence of Further Motives and Loss of Connection to Self-Preservation*: The third part of *HA* §92 describes how justice comes to be sustained by further motives as children are taught to admire and imitate just acts while the original connection to self-preservation is forgotten. Nietzsche describes this process more fully in the second volume of *HA*:

> The same actions that within primitive society were first performed with a view to common utility have later been performed by other generations from other motives: out of fear or reverence of those who demanded and recommended them, or out of habit, because one had seen them done all around one from childhood on, or out of benevolence, because their performance generally produced joy and approving faces, or out of vanity, because they were commended.

> Such actions, whose basic motive, that of utility, has been forgotten are then
> called moral actions.   (*WS*, 40)

Over generations, just acts come to be performed out of motives that have nothing
to do with self-preservation, and people forget the original connection to self-
preservation. This removes a limitation on the scope of justice, since the motive of
self-preservation makes the exercise of justice conditional on its having beneficial
consequences. Once freed of this limitation, justice may be sought even in
encounters between parties of unequal power. Socialized into seeking and admir-
ing just exchanges and forgetting their practical origins, people come to value just
exchanges independently of their value for self-preservation.

The resemblance of Nietzsche's genealogy of justice to 'English'-style geneal-
ogies that Nietzsche distances himself from in the *Genealogy* is striking. 'English'
genealogies, he remarks there, start from a hypothesis about the original function
of a certain way of valuing, before suggesting that the valuation solidifies through
habit and its function is forgotten, so that the valuation is unjustifiably extended
beyond the boundaries of its original functionality (*GM*, I, §2).

(4) *The Importance of Forgetfulness*: Nietzsche leaves us with the gnomic
claim that the world would be a lot less moral without forgetfulness, that
'doorkeeper on the threshold of the temple of human dignity'. On the reading
of this passage encouraged by the Foucauldian expectation that genealogies will be
disobliging, Nietzsche means that forgetfulness about the practical origins of
non-egoistic or moral behaviour—such as being just even when it does not
serve self-preservation—is *essential* to our exhibiting such non-egoistic behaviour:
justice is unmasked as being nothing but a means to an egoistic end, and if we
exhibit it even when it does not align with self-interest, it is only because we are
forgetful. Were we not so forgetful, we would be more consistently egoistic, and
the world would see a lot less non-egoistic or moral behaviour, i.e. less of the
behaviour regarded as expressive of 'human dignity'. On this reading, Nietzsche's
genealogy subversively and reductively identifies the product with its practical
origins: all justice, even now, is self-preservation at best and an unwarranted
product of collective amnesia at worst.

But there is also a different reading of the genealogy on which it is not reductive.
Nietzsche's conclusion that the world would be a lot less moral without forget-
fulness could mean that the emancipation from practical origins achieved by
forgetfulness was a *causally necessary* step in the development of non-egoistic
behaviour, but without being an essential constituent of the product—much as
scaffolding can be necessary to erecting a house, but where this had better not
entail that the resulting construction will crumble to pieces once the scaffolding is
removed. Nietzsche can then acknowledge that justice is now much more than it
originally was—as the genealogy itself shows, we have acquired further motives for
being just besides the one we originally had, and there is nothing in the story to

suggest that only prudential motives count as *bona fide* motives. Consequently, the genealogy need not be subversive if it allows that the new motives and valuations encouraging justice are genuine. These may precisely warrant being just even beyond the confines of self-preservation. Nietzsche's point is only that much of the non-egoistic behaviour commonly regarded as constitutive of human morality and dignity would never have arisen had we always concentrated on the motive at its practical origin. Forgetfulness about practical origins is a way of severing the intimate connection between justice and self-preservation, thereby allowing justice to develop into something potentially far removed from its practical origins.

There remains a question as to how seriously we should take the idea that the practical origins of our conceptual practices were literally *forgotten*—Nietzsche for one came to think that this was psychologically absurd (*GM*, I, §3). But the insight that the genealogist seeking to naturalize moral ideas by tracing them to practical needs must account for the fact that these ideas emancipated themselves from narrowly individualistic and prudential considerations remains a valuable one, which will occupy us further in later chapters.

5.3.2 Nietzsche's Early Genealogy of Truthfulness

Another early genealogy of Nietzsche's which fits the pragmatic genealogical model is his genealogy of truthfulness. Sketches towards such a genealogy appear as early as 1872 and reappear throughout later notebook entries, but particularly notable is the 1873 essay 'Über Wahrheit und Lüge im aussermoralischen Sinne' (translated, in the English edition I use here, as 'On Truth and Lies in a Nonmoral Sense').[7] This essay has received much attention from post-modernists for its discussion of the metaphorical nature of language and thought and the doubt it casts on our ability to achieve truth as correspondence with the world as it is in itself. But these Neo-Kantian doubts will not be my concern here, except to note that they help explain why Nietzsche is led to inquire into the origins of truthfulness. It is precisely the realization that the notion of truth as correspondence with the world as it is in itself is beset with difficulties that invites the question *why* we ever came to be so obsessed with truth. Nietzsche's genealogy is thus not so much a *how-possibly* explanation as a *why-ever* explanation. If we do not have access to the truth anyway, what is the point of truthfulness? Although the human intellect may now be thought of as a means of discovering truths, Nietzsche hypothesizes that in a bellicose 'state of nature' ('*in einem natürlichen Zustande der Dinge*'), it would primarily be used for deception (*TL*, §1). He thus rejects the Aristotelian

---

[7] For a detailed study of the genesis of 'On Truth and Lies', see Scheibenberger (2016).

premise that humans naturally seek the truth: 'Man does not by nature exist in order to know' (*eKGWB*, 1872, 19[178]). 'Deception, flattering, lying, deluding, talking behind the back', Nietzsche writes, is 'so much the rule and the law among humans that there is almost nothing which is less comprehensible than how an honest and pure drive to truth could have arisen among them' (*TL*, §1). What, then, is 'the value of this will' (*TL*, §1) to truth?

Nietzsche's early genealogy of truthfulness answer this question by showing that truthfulness has practical origins in the exigencies of social life.[8] What we call 'truths' may seem to fall short of corresponding to the 'world as it is in itself'—at least when that phrase is taken in the elusive metaphysical sense given to it by Neo-Kantians like Friedrich Lange, who so impressed Nietzsche during his studies (Blue 2016, 237); when measured by that standard, our 'truths' appear to be only 'illusions' and 'lies'.[9] But the practical demands on coexisting human beings force them to draw some contrast between descriptions of the apparent world that are misleading and dangerous and those that are less so. It is from this pressure that our concern with truth stems. It has its origin not in an epistemological contrast between truth and falsity, but in a deontological contrast between truth and lies.

Nietzsche uses a variety of related terms in speaking of truthfulness—'truthfulness', 'the will to truth', 'the love of truth', 'the pathos of truth', 'honesty', 'the drive to truth'. What unifies them is that they express human concern with the truth, expressed most basically in one's making an effort to see things as they are, undistorted by wishful thinking, lies, and deception.[10] Nietzsche sometimes uses these terms in connection with dispositions to *seek* the truth (*eKGWB*, 1872, 19 [175–7]), and sometimes in connection with dispositions to *tell* the truth (*eKGWB*, 1872, 19[207]). We can therefore distinguish between truthfulness as *truth-seeking* and truthfulness as *truth-telling*. Nietzsche's genealogy of how the dispositions of truthfulness arose involves six steps:

(1) *Entry into Society and Language*: The first is the entry into society and language. Nietzsche tells us that man, 'from boredom and necessity', wishes to 'exist socially and with the herd; therefore, he needs to make peace and strives accordingly to banish from his world at least the most flagrant *bellum omni contra omnes*' (*TL*, §1). This peace treaty

---

[8] One of the few interpreters to note this aspect of 'On Truth and Lies' is Alexander Nehamas, who writes that the essay presents the origin of truthfulness as 'profoundly practical'—it 'locates the origin of the drive for truth and knowledge in our need for social organization' (2012, 32).

[9] Later, Nietzsche became more critical towards the claim that we do not have access to the world as it really is (Clark 1990; *TI*, True World). He came to see that the idea that the True World is systematically being falsified by our constitution-laden descriptions of it incoherently presupposes a comparison with an unintelligible standard: 'The antithesis of the apparent world and the true world is reduced to the antithesis "world" and "nothing"' (*eKGWB*, 1888, 14[184]).

[10] See Nietzsche (*GS*, P, §4; *BGE*, §§9, 230).

brings in its wake something which appears to be the first step toward acquiring that puzzling truth drive: that which from now on shall count as 'truth' is established. A uniformly valid and binding designation is invented for things, and this legislation of language likewise establishes the first laws of truth. For the contrast between truth and lie arises here for the first time. The liar is a person who uses the valid designations, the words, in order to make something which is unreal appear to be real.   (*TL*, §1)

(2) *Emergence of a Prototypical Form of Truth-Seeking*: The second step is the emergence of the prototypical form of the 'will not to let oneself be deceived' (*GS*, §344), i.e. the disposition to seek out the truth and get one's beliefs right. Nietzsche argues that the will not to let oneself be deceived, which now grounds the scientific pursuit of truth, originally emerged out of a much narrower concern with the *consequences* of deception. Truth-seeking originally arises for instrumental reasons, as a means of satisfying human beings' need to avoid 'not deception itself, but rather the unpleasant, adverse consequences of certain kinds of deception'; it is 'in a similarly restricted sense that human beings now want nothing but truth: they desire the pleasant, life-preserving consequences of truth; they are indifferent to pure knowledge if it has no consequences' (*TL*, §1). Prudence and mistrust are the motives that originally drive humans to truthfulness as truth-seeking.

As Nietzsche's notebooks make clear, this includes mistrust towards oneself. An inaccurate grasp of one's needs or fears can be as harmful as deception by others. 'In dealing with what lies outside, danger and caution demand that one should be on one's guard against deception: as a psychological preconditioning for this, also in dealing with what lies within'. 'Mistrust', he writes, is 'the source of truthfulness' (*eKGWB*, 1885, 40[43]). Considerations of utility thus drive the emergence of truthfulness insofar as they drive the cultivation of a prudential disposition to seek out and acquire truths.

(3) *Emergence of a Prototypical Form of Truth-Telling*: The third step is the emergence of the prototype of what Nietzsche calls the 'will *not to deceive*' (*GS*, §344), i.e. the disposition honestly to tell what one takes to be the truth. Here the individualistic approach pursued so far runs into an obstacle: from a purely instrumental point of view, truth-telling must appear mostly unattractive, since its value largely consists in its value to others. Hence, the reasons one might give in answer to the question 'But why not deceive?' must lie in 'a completely different area' from those one might give when asked 'But why not let oneself be deceived?' (*GS*, §344).

Already in 1872, Nietzsche's solution to the problem of truth-telling's emergence is to switch from the *individual* to the *social* point of view. Though the individual has little reason to cultivate truth-telling, there is a *collective need* to do so within society as a whole. This aspect of truthfulness 'makes its appearance as a

social need' (*eKGWB*, 1872, 19[175]); 'necessity produces truthfulness as a society's means of existence' (*eKGWB*, 1872, 19[177]). Truth-telling is necessary to society's existence because social cohesion and cooperation would break down in the face of a general fear of being deceived. In one of his earliest notes on the origin of truthfulness, Nietzsche writes: 'One anticipates the unpleasant consequences of reciprocal lying. From this there arises the *duty of truth*' (*eKGWB*, 1872, 19[97]). As he puts it in *Truth and Lies*, there is 'a duty to be truthful which society imposes in order to exist' (*TL*, §1). What one has a duty to do, in particular, is to conform to linguistic convention in order to represent things as one takes them to be. If someone 'misuses fixed conventions by means of arbitrary substitutions or even reversals of names', and 'does this in a selfish and moreover harmful manner, society will cease to trust him and will thereby exclude him' (*TL*, §1). The threat of ostracism in turn gives the individual a prudential reason to tell the truth.[11]

In a later note, Nietzsche spells out the imperative that society addresses to the individual thus:

> You shall be knowable, express your inner nature by clear and constant signs—otherwise you are dangerous: and if you are evil, your ability to dissimulate is the worst thing for the herd. We despise the secret and unrecognizable.—Consequently you must consider yourself knowable, you may not be concealed from yourself, you may not believe that you change. (*eKGWB*, 1883, 24[19])

This last sentence opens up a vista on two further, and connected, thoughts. One is that '"I do not want to deceive myself" is included as a special case under the generalization "I do not want to deceive"' (*GS*, §344), because a self-deceived informant is as unhelpful as a lying one. The other is that 'the demand for truthfulness presupposes the knowability and stability of the person' (*eKGWB*, 1883, 24[19]). How so? Nietzsche's answer seems to be that truth-telling can only have practical value insofar as finding out what individuals *really* believe or desire possesses *predictive value*, and this is only the case if these beliefs and desires display a certain degree of stability. Part of the reason why others want to know what I believe and desire is that they want to rely on that information in predicting how I will behave, and they can only do that insofar as my beliefs and desires exhibit some stability. This is an application of Nietzsche's point that rendering the individual fit for coexistence in society involves '*making*' the individual 'to a certain degree necessary, uniform, like among like, regular, and accordingly predictable' (*GM*, II, §2). Moreover, the sincere expression of one's beliefs and desires *must* exhibit a minimal amount of stability over time if it is to *count* as a

---

[11] Pettit (2018, ch. 2) gives a notably similar story about how truth-telling emerges out of the fear of ostracism.

sincere expression at all. The cultivation of truthfulness as truth-telling requires the cultivation, *to a certain degree*—Nietzsche is clear that full-blown essentialism is not called for—of a belief in the relative constancy of human beings: 'it is the object of education to create in the herd member a *certain degree of belief* in the essence of man: it is *only at this point that this belief is generated*, so that "truthfulness" can then be demanded' (*eKGWB*, 1883, 24[19]). Since Nietzsche holds that it is with the help of the *Sittlichkeit der Sitte* that 'man was made truly calculable', this suggests that truth-telling already formed part of the *Sittlichkeit der Sitte*, and thus of the 'true work of man on himself for the longest part of the duration of the human race, his entire *prehistoric* work' (*GM*, II, §2).

But if it is practical demands—first at the level of the individual, and then at the level of society—that explain the emergence of truthfulness, these practical demands also impose limits on what forms of truthfulness are practically warranted: limits concerning *whom* one needs to be truthful towards and *how much* information one needs to convey. For example, it is not clear that one should be truthful towards people outside one's community: 'Within a herd, within any community, that is to say *inter pares*, the overestimation of truthfulness makes good sense', namely as a 'mutual obligation between peers!' (*eKGWB*, 1885, 40 [43]). He spells this out later: 'One says what one thinks, one is "truthful", only under certain conditions: namely, that one is understood (*inter pares*), and understood charitably (once again *inter pares*). One conceals oneself in presence of the unfamiliar' (*eKGWB*, 1886, 7[6]). The social imperative to be truthful is originally restricted to encounters among community members (as Don Corleone puts it in *The Godfather*, never let anyone outside the family know what you are thinking). Even then, there is still a question about how far one must disclose one's thoughts to satisfy the requirements of truthfulness. Cleary, Nietzsche did not think that truthfulness implies complete disclosure: 'the demand that one should *denude oneself* with every word one says is a piece of naiveté' (*eKGWB*, 1886, 7[6]).

Originally, then, it stands with truthfulness much as it stands with deceit: it has limited application, and what value it has is the practical value it derives from its consequences. Truthfulness is useful from the point of view of the individual as a means of controlling the world, of avoiding deception, of communicating effectively, and of avoiding exclusion from society; and it is useful from the point of view of society as a means of securing social cohesion, cooperation, and stability. In each case, a calculus of utility explains, sustains, and justifies truthfulness.

Yet our attitude towards truth is not conditional on its possessing instrumental value for us. We are, as Nietzsche puts it, *unconditionally* truthful. How did such an unconditional drive to truth emerge?

> Where might science get the unconditional belief or conviction on which it rests, that truth is more important than anything else, than every other conviction? Precisely this conviction could never have originated if truth *and* untruth had

constantly made it clear that they were both useful, as they are. So, the faith in science, which after all undeniably exists, cannot owe its origin to such a calculus of utility; rather, it must have originated *in spite of* the fact that the disutility and dangerousness of the 'will to truth' or 'truth at any price' is proved to it constantly.   (*GS*, §344)

Nietzsche stresses the disutility of unconditional truthfulness, and one might think he overdoes it a little. Truthfulness that is unconstrained by considerations of utility—and thus unconditional in the weak sense of not being directly subservient to further aims—has its uses. It is, for instance, one argument for pure as opposed to applied research that the scientific enterprise proves more useful in the long run if it is not always guided by the potential for useful application.

But the attitude that takes truth to be 'more important than anything else' and strives for truth 'at any price' (*GS*, §344) is unconditional in a further and stronger sense: it considers truth to be not merely an end in itself, but an end to be realized *under any conditions*. We can thus distinguish three forms of truthfulness:

   (i) *Conditional truthfulness*: truth is valued only instrumentally, as a means to an end.
  (ii) *Unconditional truthfulness as a* pro tanto *reason*: truth is valued in itself as providing a *pro tanto* reason for action.
 (iii) *Unconditional truthfulness as an overriding reason*: truth is valued in itself as providing an overriding reason for action—a reason that, even all things considered, trumps every other reason.

Unconditional truthfulness as an overriding reason is what Nietzsche gives voice to when he writes: '*Nothing* is *more* necessary than truth; and in relation to it, everything else has only secondary value' (*GS*, §344). Such an attitude, he thinks, cannot be vindicated by considerations of utility, because 'there is no pre-established harmony between the furthering of truth and the well-being of humanity' (*HA*, §517). Nietzsche was fond of quoting a line of Byron's on this point: 'The tree of knowledge is not that of life' (*HA*, §109). How, then, could such an unconditional and overriding form of truthfulness have arisen?

(4) *Forgetting the Original Function of Truthfulness*: The fourth step in Nietzsche's genealogy is that truthfulness' original function is forgotten. As we saw, forgetfulness about practical origins is a tendency that Nietzsche also invokes in his account of the origins of justice, and he seems to regard it as a general human tendency. People are truthful 'in accordance with centuries-old habits' (*TL*, §1), and forget *why* it was originally cultivated. This loss of the conscious connection to individual and social needs clears the ground for truthfulness to emancipate itself from its practical origins.

(5) *Moralization of Truthfulness*: The fifth step is the development of truthfulness from an instrumental into an *independently motivating* reason for action. One is truthful not because doing so has beneficial consequences, but just because that is the kind of action it is. It is thanks to the fact that the prudential motives to truthfulness have moved out of sight with the forgetting of truthfulness' original function that the socially imposed duty to be truthful can generate what Nietzsche calls *moral* motives:

> precisely because of this unconsciousness, precisely because of this forgetting, one arrives at the feeling of truth. From the feeling that one is obliged to designate one thing as 'red', another as 'cold', and a third as 'mute', there arises a moral impulse in regard to truth; from its opposite, the liar whom no one trusts and all exclude, human beings demonstrate to themselves just how honourable… truth is.   (*TL*, §1)

The habit of truthfulness, heretofore understood only as a *prudentially motivated* disposition, is given a moral gloss and becomes a virtue: 'our habits become virtues', Nietzsche suggests, because we 'include inviolability within the concept' of the behavioural patterns we are in the habit of engaging in—'because we consider their inviolability to be more important than our own particular welfare' (*eKGWB*, 1872, 19[185]). This is the 'recoining of habit as virtue, of *Sitte* as *Sittlichkeit*', which Nietzsche dubs a 'fine old—age-old—piece of counterfeiting' (*eKGWB*, 1882, 3[1]). The Hegelian phrase *Sittlichkeit der Sitte*, an enduring element in Nietzsche's later thought, points to the normative force of habits and the weight of precedent, which comes into play whenever the fact that particular patterns of behaviour have been unbroken in the past itself becomes a reason not to break them: '*Sitte* represents the experiences of men of earlier times as to what they supposed useful and harmful—but *the feeling for the Sitte* (*Sittlichkeit*) applies, not to these experiences as such, but to the age, the sanctity, the indiscussability of the *Sitte*. And so this feeling hinders the acquisition of new experiences and the correction of *Sitten*' (*D*, §19).[12] Useful dispositions arise because they are useful; but they are held in place by ties that are less conditional than those of prudence: those of moral feeling. It is when habits become anchored in feelings of inviolability that with truthfulness we '*stand on moral ground*' (*GS*, §344).

(6) *Metaphorical Extension of Truthfulness' Domain of Application*: The sixth and final step in Nietzsche's genealogy is the *metaphorical extension* of truthfulness'

---

[12] *Sitte* and *Sittlichkeit* are often translated as 'custom' and 'morality', but this obscures the way in which Nietzsche takes the etymological connection between them as a guide to their more substantive connection: *Sitte* breeds *Sittlichkeit* which in turn stabilizes *Sitte*.

domain of application. It is the coupling of the moral notion of truthfulness with what Nietzsche considers to be a 'fundamental human drive'—the 'drive to form metaphors' (*TL*, §2). It is this synthesis that produces the unconditional and disinterested drive to truth or knowledge:

> Under certain circumstances, necessity produces truthfulness as a society's means of existence. Through frequent practice, this drive is reinforced and is now, by means of metastasis, unjustifiably transferred. It becomes an inclination in itself. A quality [i.e. truthfulness] develops out of a practice [developed] for specific cases.—Now we have the drive to knowledge. This generalization takes place by means of the intervening *concept*. This quality begins with a *false* judgment:—to be true [i.e. truthful] means to be true [i.e. truthful] *always*. From this arises the inclination to live without lies: elimination of all illusions....Two qualities, each required for a different purpose, have produced the inclination to truth—*truthfulness*—and *metaphor*. Thus the intellectual drive is produced by an aesthetically generalized moral phenomenon.
>
> (*eKGWB*, 1872, 19[177–8])

Having been brought into existence by individual fear of deception, augmented by a social imperative not to deceive others and transformed by forgetfulness and the force of habit into a moral notion, the concept of truthfulness acquires a significance that is independent of individual or social exigencies. This in turn leads to its being applied to circumstances beyond those that drove its emergence. Having made its 'appearance as a social need', 'by means of a metastasis, it is then applied to everything, where it is not required' (*eKGWB*, 1872, 19[175]). Nietzsche's talk of 'metastasis' and 'metaphor' indicates that a *transfer* has taken place from the sphere of application in which truthfulness originally had its home and was instrumentally justified to spheres of application where it is no longer instrumentally justified (hence the transfer's description as 'unjustifiable'). It is an *aesthetic* generalization because 'between two absolutely different spheres...there is no causality, no correctness, no expression, but at most an *aesthetic* way of relating' (*TL*, §1)—that is, the generalization is not *rationally intelligible*, but is driven only by associative or analogical thinking.

   Among the contingent extensions of truthfulness, Nietzsche suggests, was the transfer of truthfulness from the *social* to the *natural* sphere. It came to be expected not only that other people would be truthful towards oneself, but also that nature would follow suit: one would be granted access not only to the real opinions of other people, but also to the world as it really is—when 'man sets up truthfulness as a law for himself, he also believes in the truthfulness of nature towards him' (*eKGWB*, 1872, 19[207]). He 'transfers his inclination to truth to the world and believes that the world must in turn be true towards him' (*eKGWB*, 1872, 19[177]). This rather blankly psychological assertion finally bridges the gap

to the question we started out from—of why, as the Nietzsche of the early 1870s still believed, humans self-importantly imagine themselves to be discovering the world as it really is when they have access only to illusions. We have come full circle.

With this reconstruction of Nietzsche's genealogy of truthfulness in place, we are in a position to draw out its resemblance to 'English'-style genealogy, which is even more pronounced than that of his genealogy of justice. Following just the pattern he outlines in *GM*, I, §2, Nietzsche's genealogy of truthfulness starts out from a hypothesis about the original function of truthfulness, and then suggests that while it solidifies through habit, its function is forgotten, so that it is erroneously extended beyond its original domain of application. The puzzling truth drive is thus 'an extension or a solidification of a way of thinking and acting which was necessary in certain cases' (*eKGWB*, 1872, 19[178]).

This licenses the conclusion that at least two of Nietzsche's own genealogies were in the 'English' style, explaining the emergence of practices in terms of their original functionality and invoking forgetfulness as the mechanism by which they could subsequently outgrow a merely functional understanding and develop a life of their own.

5.4 Hypertrophy: Taking a Good Thing Too Far

How are Nietzsche's early genealogies supposed to bear on the value of their object? Contrary to Nietzsche's reputation as a genealogical debunker, these genealogies clearly possess vindicatory aspects. Much like Hume's genealogies, they *vindicate* justice and truthfulness against the suspicion that they might be metaphysical *ignes fatui*, explaining how they could naturally have arisen and displaying their practical value in satisfying individual and social needs. In this respect, the genealogies are representative of Nietzsche's treatment of virtues in general: 'All virtues', he remarks already in 1872, 'arise from pressing needs' (*eKGWB*, 1872, 19[175]). He shares Hume's methodological assumption—further encouraged by the Darwinian revolution—that what we deem virtuous is non-coincidentally related to what helps us to live, and a pragmatic and naturalistic genealogical explanation that exploits this connection will in the first instance prove vindicatory. Fifteen years later, Nietzsche still conceives of virtues in practical terms as contributing to the effective operation of society: 'I attempt an *economic* justification of virtue.—The task is to make man as useful as possible and to approximate him, as far as possible, to an infallible machine: to this end he must be equipped with the virtues of the machine'; and because the states in which he is useful are not those he would be drawn to out of self-interest, 'he must learn to experience the states in which he works in a mechanically useful way as the supremely valuable states'—they must be 'enveloped in a higher charm' (*eKGWB*,

1887, 10[11]).[13] As a phrase in the *Nachlass* has it, there is a 'point' to 'society's *myopic* perspective with regard to usefulness' (*eKGWB*, 1888, 12[1]). In its functionalist derivation of non-functionalist ways of thinking, these and similar passages (*eKGWB*, 1887, 10[8, 10, 57]) prefigure the idea that we later find fleshed out in Williams's genealogy: that intrinsic values—charms higher than the allure of the instrumentally valuable—have their uses. It can be more useful in the long run to cultivate a mindset that focuses myopically on the immediate value of doing things for their own sake rather than on their long-term usefulness. This initially vindicatory account of virtues as 'economically justified' does not prevent Nietzsche from going on to criticize how these virtues stifle the creativity of exceptional individuals. On the contrary, it provides a foundation for that critique by explaining why the stifling forces are there in the first place and why they really would be stifling.

Nietzsche thus differs from Hume in his sensitivity to the fact that while pragmatic genealogies are in the first instance vindicatory, this does not entail that they are vindicatory all things considered, i.e. once one takes into account the historical and social elaboration of conceptual practices and the differences in concept-users' needs and capacities. Nietzsche's genealogy of truthfulness exemplifies this: although it shows that truthfulness originally responds to practical exigencies, even this early genealogy already identifies pragmatically unmotivated elaborations of truthfulness. As both the value and the scope of truthfulness are inflated to the point where it is demanded always and 'at any price' (*GS*, P, §344; *eKGWB*, 1872, 19[97]), it becomes a 'hypertrophic virtue' (*eKGWB*, 1873, 30[2]): a good thing has been taken too far, because its sphere of application or its normative weight goes beyond what makes practical sense for us. Truthfulness should not be pursued under *any* circumstances or at *any* price. The hypertrophic form of truthfulness expressed in the motto '*fiat veritas pereat vita*' (*UM*, II, §4), let truth prevail though life perish, has clearly turned an originally life-promoting idea into a life-denying one.[14]

The problematic historical elaborations of the prototype of truthfulness come to the fore in Nietzsche's later work, particularly in the *Genealogy*. While his early genealogy explained the emergence of truthfulness as an overriding reason by appeal to a metaphorical drive and our natural forgetfulness, this explanation in terms of natural propensities is abandoned in the *Genealogy* in favour of an explanation in terms of asceticism's nefarious influence. In what can be read as a revision of the later stages of his 1870s genealogy of truthfulness, Nietzsche writes that if a moralized distinction between truthfulness and lying is available in a socially unequal society, as the earlier parts of his genealogy suggest it would be, that distinction will be harnessed by the upper caste to articulate their superiority.

---

[13] See also Nietzsche (*TI*, Skirmishes, §29).
[14] See Nietzsche (*GM*, III, §7; *eKGWB*, 1873, 29[8]).

They will describe themselves as 'the truthful ones', as 'distinct from the *lying* common man' (*GM*, I, §5).[15] It is this development—rather than some propensity to overgeneralization—that entrains the transformation of truthfulness into something potentially life-denying. For while the 'priestly' types in the upper caste flourish best by cultivating and disseminating ascetic values of self-restraint and self-abnegation, what is life-promoting for priestly types is not necessarily life-promoting for other types;[16] and as Nietzsche argues in the Third Treatise of the *Genealogy*, the influence of priestly asceticism has rendered truthfulness dangerously hypertrophic, so that he must take his fellow scholars to task for practising a form of self-restraint and self-abnegation in the name of truth and objectivity that rivals the most life-denying asceticism.[17]

This is a condemnation of a specific elaboration of truthfulness rather than a wholesale rejection of it. Nothing in the *Genealogy* touches on Nietzsche's earlier insights into the practical value of more modest forms of truthfulness, suggesting that his later critique coexists with—indeed builds on—his earlier pragmatic vindication. What the earlier genealogy does not yield are reasons for regarding truthfulness as an overriding reason for action that licenses the attitude of *fiat veritas, pereat vita*, and it is this hypertrophic form of truthfulness that Nietzsche warns against in the *Genealogy*.

Reading Nietzsche's later critique of truthfulness as rooted in an earlier pragmatic vindication of it helps explain why truthfulness survives the revaluation which, on Nietzsche's view, ushers in Slave morality. The 'fear-inspiring consistency' with which the 'aristocratic value equation (good = noble = powerful = beautiful = happy = beloved of God)' was inverted to suggest that the good and blessed in God are the 'miserable', 'poor', 'powerless', 'suffering', and 'ugly' (*GM*, I, §7) did not go so far as to invert the rank order between truthfulness and lying.[18] Nietzsche does not account for this exemption, but our reconstruction of his earlier thoughts on the matter makes it clear that truthfulness is held in place by *practical exigencies*. Which particular form truthfulness takes is a matter of historical elaboration, but Nietzsche's genealogy has shown that *some* form of truthfulness is necessary to the satisfaction of both individual and social needs. The prototypical form of truthfulness he has outlined is part of the minimal ethical consciousness that renders social coexistence possible in the first place.

This also helps explain why some form of truthfulness survives a second revaluation of values, namely that envisaged by Nietzsche himself. Nietzsche

---

[15] Shapin (1994) lends historical support to this idea.

[16] Nietzsche, following Lecky (1869, 156–7) and other authors of his time, distinguishes different character types in his theorizing. See Leiter (1998; 2015, 6, 54, 72, 128, 252) for an influential discussion.

[17] See Gemes (1992, 2006). Daston and Galison (1992, 83, 121–2) describe the asceticism involved in nineteenth-century science.

[18] For a historical study of the rise of Christian morality that offers a lot of grist to Nietzsche's mill, see Holland (2019).

praises, and remains committed to, a stringent sort of truthfulness—though not as stringent and certainly not as unquestioning as the attitude of *fiat veritas, pereat vita*.[19] Nietzsche's early thought tells us little about why he later valued truthfulness *to the extent* that he did—truthfulness, he later notes, requires 'greatness of soul' (*AC*, §50), and he takes how much truth a spirit can endure to be 'the real measure of value' (*EH*, P, §3). But it does tell us something about why the later Nietzsche continued to hold on to truthfulness in *some* form. David Owen (2007) has offered one explanation along these lines, suggesting that Nietzsche was implicitly concerned with reflective stability: 'unless truthfulness is an intrinsic value for Nietzsche, his project of re-evaluation will not possess the right kind of reflective stability' (2007, 70). It is true that Nietzsche would be sawing off the branch that supports his genealogical inquiry if he were to renounce truthfulness altogether. Yet in light of the above, we can add a second rationale for Nietzsche's enduring commitment to some form of truthfulness: he was well aware that the need for truthfulness forms a corollary of the demand for *social* as much as for reflective stability. His earlier derivation of non-hypertrophic truthfulness from basic individual and social needs had made it clear to him that truthfulness was something we could not do without.

5.5 Thinking Historically

We can now see that Nietzsche's rebuke of the 'English genealogists' contains a non-negligible dose of self-criticism: abandoning his earlier explanation in terms of natural propensities towards metaphorical extensions and forgetfulness, the mature Nietzsche recognizes the explanatory importance of *local* needs. Needs can be local both in the sense of being specific to certain character types and in the sense of being specific to certain cultures and epochs, because they are needs that were 'implanted' (*eKGWB*, 1886, 5[71]) in us by socio-historical developments. In describing the elaboration of truthfulness in his own socio-historical situation, for example, Nietzsche invokes the historical impact of Christianity and the local needs of priestly types. But does this recognition of the importance of history amount to an *amendment* or an *abandonment* of his earlier method of pragmatic genealogy?

Some have felt that recognizing the historicity of our ideas precludes a functionalist treatment of them. Alexander Prescott-Couch, for example, has argued that acknowledging the role of history in shaping our ideas forecloses the possibility of explaining them in terms of their function, because that history suggests that there no longer is such a thing as *the* function served by an idea:

---

[19] See Owen (2003, 2007), Reginster (2013), Jenkins (2016).

When trying to assess the value of the Catholic Church, it would be misguided to proceed by first attempting to determine its function and then evaluating that function. This procedure would be misguided because there is no one such function.... One crucial purpose of genealogical investigation into morality's origins is to demonstrate that approaches to assessing morality by first inquiring into its functional role in social life are similarly misguided.

(Prescott-Couch 2015b, 107)[20]

It is certainly true that Nietzsche is sensitive to the fact that needs and functions are in principle subject to change. Conceptual practices get repurposed. We cannot simply assume that we can infer from original to current function or *vice versa* (*GM*, II, §12).[21] This is why Nietzsche criticizes the English genealogists for ignoring vast stretches of history when they equate current with original function. It is also true that Nietzsche contrasts 'English hypothesizing *into the blue*' with genealogizing that looks to the 'real *history of morality*', the '*grey*', that which 'can be documented, which can really be ascertained, which has really existed' (*GM*, P, §7).

But Nietzsche's problem with English genealogists is not that they ask the wrong question or that they formulate hypotheses about functions at all. His fire is directed at hypothesizing that completely ignores and remains unconstrained by history:

I had no doubt that [Dr. Rée] would be pushed by the very nature of his questions to a more correct method of attaining answers.... My wish...was to turn so sharp and disinterested an eye in a better direction, the direction of the real *history of morality* and to warn him while there was still time against such English hypothesizing *into the blue*. It is of course obvious which colour must be a hundred times more important to a genealogist of morality than blue: namely *grey*, which is to say, that which can be documented, which can really be ascertained, which has really existed, in short, the very long, difficult-to-decipher hieroglyphic writing of the human moral past! *This* was unknown to Dr. Rée; but he had read Darwin:—and thus in his hypothesizing we have, in a manner that is at least entertaining, the Darwinian beast politely joining hands with the most modern, unassuming moral milquetoast who 'no longer bites'....

(*GM*, P, §7)

A reading of the later Nietzsche as rejecting functional hypothesizing altogether misses three noteworthy points in this passage. First, Nietzsche sees the need for a

---

[20] See also Prescott-Couch (2014, 2015a).
[21] Nietzsche also highlights further distinctions in this passage, such as that between the causes of the first emergence of something and the reasons for its retention.

turn towards history as arising out of Rée's *own* question, which is a question about practical origins. Second, the emphases suggest that Nietzsche takes issue not with hypothesizing *per se*, but with unconstrained hypothesizing *into the blue*; such free roaming of the imagination is to be constrained by looking *towards* real history—*informing* rather than replacing hypothesizing by real history. Third, the passage is generally comparative and corrective rather than contrastive: it is a matter of giving the same eye a *better* direction; Rée's method would be *more correct* if it displayed greater awareness of real history; the grey of documented facts should be *more important* to the genealogist than the blue of the merely hypothesized. This is neither a wholesale rejection of Rée's question and method nor of hypothesizing. It is the complaint that Rée ignores the history that lies between the Darwinian beast and the modern milquetoast—they are lumped together as fundamentally similar, while on Nietzsche's view they could not be more different, and there is a complex story to be told about how we got from one to the other. To take this story into account, however, is not to jettison Rée's functionalist hypothesizing wholesale, but rather to acknowledge that functional hypothesizing must be 'hardened and sharpened under the hammer-blow of historical knowledge' (*HA*, I, §37). What Nietzsche says and does in the *Genealogy* suggests that he advocates *historically informed functional hypothesizing*. In the terminology of Chapter 3, he advocates pragmatic genealogy tailored to socio-historical situations.

This is evidenced in the *Genealogy* by the fact that needs are increasingly localized in both epochs and character types. For example, Nietzsche distinguishes the needs of the strong from the needs of the weak (*GM*, I, §4), or the needs of the knightly-aristocrats from those of the priestly-nobles (*GM*, I, §6). The historical fact that a new 'manner of valuation branches off from the knightly-aristocratic' (*GM*, I, §7) is incorporated into the genealogical narrative and modelled as a reversal of values driven by *ressentiment* (*GM*, I, §§7–10). The contest between various instantiations of these two manners of valuations ('good and bad' and 'good and evil') is then—however roughly—situated in history: it pits the Roman ideal against the Judaeo-Christian ideal, the 'privilege of the few' against the 'privilege of the masses', the Renaissance against the Reformation, the French of the *Âge Classique* against the French of the *Revolution*, and Napoleon against Europe (*GM*, I, §16). Such vague references might make sense as markers or placeholders for the historical developments that his dynamic models are being tailored to accommodate—but they would be absurdly noncommittal if understood as attempts at documentary history.

Even when interpreted as dynamic models, however, these sketches towards a genealogy of morality can feel overly compressed and vague when compared to the more focused, piecemeal genealogies of values exemplified by his earlier treatments of justice and truthfulness. For his grand genealogical synthesis, Nietzsche employs what might be called the *carriage return method* of narration:

he separates in his genealogical narrative what goes together in reality, sequentially recounting developments that actually run in parallel. Each treatise focuses on the fusion, development and branching off of different strands which all come together in the '*moral in a narrow sense*' (*BGE*, §32)—the socio-historically localized constellation of values associated, in particular, with Christianity. Nietzsche goes back and forth between these strands, pursuing one and then returning to pick up another, just as the carriage return goes back and forth on a typewriter. This method grows naturally out of an earlier insight that Nietzsche formulated in connection with his analysis of saintliness and asceticism: 'almost everywhere, within the physical world as well as in the moral, the supposedly miraculous has successfully been traced back to the complex, to the multiply caused. Let us therefore venture first to isolate individual drives in the soul of the saint and ascetic and then conclude by thinking of them entwined together' (*HA*, §136). The *Genealogy* proceeds along similar lines. In a postcard to Franz Overbeck, his colleague and housemate in Basel who specialized in the secular origins of Christianity,[22] Nietzsche explains that 'it was necessary, for the sake of clarity, to isolate artificially the different roots of the complex structure that is called morality' (*eKGWB*, BVN 1888, 971). The book indeed isolates such roots in the way in which the pathos of distance, the instincts of freedom, and the preservation and healing instincts express themselves in certain types under certain circumstances, and then sketches their elaboration and combination issuing in the constellation we actually find. The result has proven to be a powerful and enduringly fascinating book. But as far as exemplifying a genealogical *method* is concerned, Nietzsche's ambition to genealogize the whole of morality in twenty days (the time he said elapsed between its conception and completion) and fewer than 50,000 words leaves us with a work which, ironically, is far more methodologically ambiguous than his earlier genealogies.[23] Despite—and indeed also because of—the methodological reflections at the beginning and midway through the book, Nietzsche's later genealogical method can seem rather underdetermined by its application.

As a result, there is plenty of room for a reading on which Nietzsche is not abandoning pragmatic genealogy, but improving on it by its own lights, encouraging it to *look* to history without *giving way* to history. The *Genealogy* remains, after all, a far cry from the patiently documentary history that Foucault (1971) envisages. But it does mark an advance over Nietzsche's earlier genealogies in its insistence that genealogists should *think historically*—an injunction that must appear pleonastic if genealogy is equated with history, but which is very much

---

[22] For a discussion of how Overbeck's research on early Christianity relates to Nietzsche's genealogical method, see Sommer (2003).

[23] See Sommer (2019, 3–9) for a detailed account of the genesis of the text, which does seem to have been sent off to the publisher less than two months after its conception.

to the point if addressed to pragmatic genealogists. To achieve what it sets out to achieve—to understand what our values *do* for us—functional hypothesizing should be informed and constrained by history.

On this reading, Nietzsche's call to heed the grey is an *internal* critique of pragmatic genealogy. This makes sense in particular if he is himself no stranger to English genealogy, and one of the findings of this chapter has indeed been that Nietzsche himself comes to the grey from the blue—he develops his historically informed genealogical method out of his own attempts at ahistorical functional hypothesis-mongering, and even a cursory reading of the *Genealogy* suggests that he does not abandon the imagined for the documented, but rather mixes the two in a greyish blue.

It is perhaps more easily accepted that a genealogy combines the vindicatory and the subversive than that it combines the blue and the grey, the imagined and the documented. But drawing on imagined situations need not make a genealogy *purely* hypothetical in the way that justificatory fictions about nasty, brutish, and short lives in the state of nature perhaps are. Instead of classifying genealogies according to whether they involve imagined or documented elements, we can distinguish genealogies from each other and from more orthodox historiography in terms of the *questions* to which they form an answer. In Nietzsche's case, the question is: What is the value of a given way of valuing? As David Owen has emphasized (2007, 142–3), it is unsurprising that a genealogy seeking to answer that question will be tailored to demands of salience, perspicuity, and persuasiveness that are quite different from the demands on answers to more Hobbesian or Foucauldian questions. Such demands might be met by offering a model constructed out of a range of resources—not only ascertained facts, but also hypotheses about how the genealogized item relates to human needs and psychology. And perhaps these relations are best represented using simplifications and distortions. In combining the documented and the imagined in this way, genealogy is no different from art, whose creative liberties can render it more truthful, or science, whose idealizing and distorting assumptions can have the same effect. Like art and science, genealogy reminds us that there is such a thing as truthful imagination.

The conclusion we reach is that Nietzsche can be brought into the fold of pragmatic genealogy, and that consequently, pragmatic genealogy can claim a genuinely Nietzschean pedigree. His distinctive contributions to the method of pragmatic genealogy are that he highlights the possibility of the originally functional growing into something hypertrophic and dysfunctional, and that he emphasizes the need for pragmatic genealogists to think historically. Functions and needs change, not only from one epoch to the next, but also from one type of individual to the other. In emphasizing this, Nietzsche helps us see why paradigm-based explanation and similarly ahistorical approaches must be insufficient to understand the accumulated multi-purposiveness and perspectival

nature of the value of our conceptual practices. To master these complexities, we need historically informed pragmatic genealogy.

It follows that some later attempts at 'English' genealogy turn out to have a more genuinely Nietzschean pedigree than previously supposed. One such attempt, to which we now turn, is Edward Craig's *Knowledge and the State of Nature* (1990).[24]

---

[24] In *Knowledge and the State of Nature*, Craig mentions neither Nietzsche nor genealogy; but see Craig (2007).

6

Loosening the Need-Concept Tie

Edward Craig

What drives people to develop the concept of knowledge? There seems to be a word for 'know' in every language, which suggests that it answers to highly general needs of human life.[1] But what are those needs? This is the question that leads Edward Craig (1990) to offer a genealogy of the concept of knowledge. He takes his cue from Williams (1973a, 146), who pointed out that philosophers have traditionally suffered from a *déformation professionnelle* in thinking about knowledge: they have considered it from the perspective of the *examiner* who assesses whether someone knows something already known to the examiner ('Does A know that *p*?'). Yet as Williams points out, the examiner's situation is a rather peculiar one. The more natural home for the concept of knowledge is the situation of the *inquirer* who seeks to identify someone who knows what the inquirer does not yet know ('Who knows whether *p*?'). Modelling the inquirer's situation in the state of nature, Craig's genealogy describes how the need for true beliefs about the immediate environment would lead the inquirer to develop a primitive, proto-typical form of the concept of knowledge as a tool by which to identify what, given the individual's needs and capacities, would be good informants. It then describes how the concept gradually loses this indexing to subjective needs and capacities to become more like our concept of knowledge.

Craig's pragmatic genealogy is the first book-length investigation of its kind that is entirely devoted to a single concept, and it is carefully focused on the *concept* of knowledge rather than on knowledge itself. *Pace* Michael Hannon, who thinks 'it makes little sense to distinguish the attempt to become clearer about our concept of knowledge from the attempt to become clearer about knowledge as such' (2019, 32), this distinction is crucial to the method of pragmatic genealogy. A pragmatic genealogy of knowledge itself would be quite a different affair, leading us to ask why a creature would need to *have* knowledge *about* its environment rather than why it would need to become sensitive to the *presence of* knowledge *in* that environment.

Perhaps counterintuitively, a pragmatic genealogy of the concept of knowledge can remain instructive even if we take knowledge itself as given. We can

---

[1] See Wierzbicka (2018).

simultaneously illustrate this important methodological point and render Craig's genealogical approach more attractive by examining its alleged incompatibility with the currently more influential approach of 'knowledge-first' epistemology (KFE) propounded by Timothy Williamson and others. KFE involves two claims: (a) knowledge is the most general factive stative attitude,[2] an attitude one necessarily bears to a truth if one bears any other factive stative attitude to it; and (b) the concept of knowledge is an unanalysable primitive that should act as a basis for explaining related concepts like *believing*, *seeing* and *remembering*.[3]

There are three reasons why one might think Craig's genealogy incompatible with KFE.[4] The first is pointed out by Williamson himself. Though he applauds Craig's dismissal of the traditional programme of analysing the concept of knowledge, Williamson still considers Craig's project marred by its failure to acknowledge that the need for knowledge is prior to the need for true beliefs:

> [Craig's project] remains too close to the traditional programme, for it takes as its starting point our need for true beliefs about our environment . . . , as though this were somehow more basic than our need for knowledge of our environment. It is no reply that believing truly is as useful as knowing, for it is agreed that the starting point should be more specific than 'useful mental state'; why should it be specific in the manner of 'believing truly' rather than in that of 'knowing'?
>
> (Williamson 2000, 31n3)

Call this the *wrong-starting-point* problem.

The second reason is that Craig's original prototype of the concept of knowledge—*proto-knowledge*, as we might call it following Martin Kusch (2009b)—is a concept tracking 'proto-knowers' or good informants, and as Miranda Fricker highlights, Craig 'tends to describe the good informant as someone recognizable as having a true belief' (2007, 144n17). But if a proto-knower is someone who, among other things, must have a true belief that *p*, this puts the concept of belief at the heart of the prototype of the concept of knowledge; and if this prototype is still part of our concept of knowledge, as Craig claims (2007, 191), this conflicts with one of the central tenets of KFE—that the concept of knowledge does precisely *not* involve the concept of belief. The problem, then, is that if Craig defines the good informant as someone who truly believes that *p*, this ultimately puts the concept of belief at the core of the concept of knowledge, 'and therefore depicts belief as prior

---

[2] An attitude being *factive* if it is one that can only be borne to a truth, and *stative* if that attitude is a state.

[3] Williamson (2000, 33, 44, 185).

[4] See also Hannon (2019, 20–2) for a comparison of knowledge-first epistemology and Craigean function-first epistemology which emphasizes the advantages of the latter over the former.

to knowledge, so that knowledge is conceived as true belief plus a bit' (Fricker 2007, 144n17).[5] Call this the *believing-as-the-core-of-proto-knowing* problem.

Third, and perhaps most importantly, KFE insists that the concept of knowledge is primitive and explanatorily fundamental (Williamson 2000, 185); but Craig's genealogy seeks to explain the concept of knowledge in terms of something more primitive—in particular, subjective needs for a concept that is more primitive than the concept of knowledge. This also seems incompatible with KFE, since, if the concept of knowledge is indeed fundamental in the way Williamson proposes, 'there is little space for any genealogy' (Kusch 2011, 12).[6] Call this the *no-room-for-genealogy* problem.

My first aim in this chapter is to show, against these three objections, that Craig's genealogy is compatible with KFE and even lends limited succour to it. Because Craig undertakes to explain the *prima facie* plausibility of extant conceptual analyses of knowledge, he remains overly beholden to the terms in which they are given—in particular, to their preoccupation with true belief. But if Craig's genealogy is freed from the trappings of past proposals, it can dispense with the notion of true belief, reach its conclusions all the same, and even lend succour to KFE. In showing how this can be so, I vindicate an otherwise enigmatic suggestion made by Williams. Williams took Craig's book to *support* KFE, because what it really indicated, on Williams's view, was that knowledge is prior to belief.[7]

My second aim is to bring out two respects in which, beyond what it tells us about the concept of knowledge, Craig's genealogy empowers the methodological framework it instantiates by identifying practical pressures driving what I call the *de-instrumentalization* of concepts—the process whereby concepts are driven to shed the traces of their practical origins. Craig's distinctive contribution to the pragmatic genealogical method is thus to loosen the need-concept tie: he demonstrates that far from reducing our mental life to instrumental thought, pragmatic genealogy not only accommodates but in fact predicts the de-instrumentalization of concepts.

My third aim, finally, is to show how Craig harnesses pragmatic genealogy's ability to help us assess and synthesize competing accounts of concepts. In coming to understand what concepts would serve us well, we gain a sense of what our concepts are likely to be and how competing accounts of them relate to each other. As we shall see, this provides a valuable way to resolve philosophical stand-offs.

---

[5] Fricker (2007, 144n17) goes on to remark that it would in principle be open to Craig to avoid this problem. I argue for the stronger claim that this problem in fact does not arise in the first place.

[6] Kusch's own position is that this problem can be overcome: see Kusch (2009b, 90) and Kusch and McKenna (2020, 1068–70).

[7] For the claim that Williams took Craig's book to indicate that knowledge is prior to belief, see Fricker (2007, 113–14n9). Fricker herself endorses a reading on which Craig can in principle remain agnostic on the priority question (2007, 144n17).

6.1 Fictional Starting Points

Let us begin with the *wrong-starting-point* problem: Craig starts from the thought that 'human beings need true beliefs' (1990, 11) rather than from the thought that human beings need knowledge. Does this vitiate the entire project from the perspective of KFE? I think not, and elaborating why not will take us through several important methodological considerations bearing on the choice of starting points for pragmatic genealogies.

The essential first move in Craig's genealogy is not its focus on the need for true beliefs, but its focus on the practical problem one faces in the situation of the inquirer. You will be in the situation of the inquirer if you are ignorant as to whether p and need to find out, if you are to be successful in navigating your environment and satisfying your most basic needs, whether p—whether, for example, the bear went into the cave. This means that we can in principle preserve neutrality between belief-first and knowledge-first accounts without prejudice to Craig's account: we can say that human beings need truths about their environment, leaving it open whether the attitude they must bear to these truths is one of believing or knowing. Inquirers simply face a problem of the form: how to come by the truth as to whether p?

Arguably, however, careful scrutiny of the inquirer's situation and its practical challenges already leads one to the conclusion that what the inquirer must seek is really knowledge rather than true belief. Williams gives an argument to this effect (whose form and spirit foreshadows Craig's pragmatic genealogy) in his 1978 book on Descartes (2005b, 23–32). The argument turns on taking seriously the idea that the inquirer's situation is a practical situation: instead of considering inquirers as disembodied intelligences and asking, at an entirely abstract level, what their respective attitudes to truths, beliefs, or knowledge should be, we consider what form the search for truth takes in a concrete practical context. For human inquirers, this context will notably be marked by the contingent but no less important fact that they operate in a non-magical world where true beliefs do not just pop up when they wish for them, but have to be sought out and identified as such. This imposes important constraints on what the search for truth can look like in practice. *Even if* we assume that our inquirers initially want truth in the form of true beliefs rather than knowledge, these constraints entail that seeking truths and seeking knowledge cannot ultimately be two separate enterprises, because once we consider how the search for true beliefs would have to be *operationalized*, we realize that in our world, that search must take the form of a search for knowledge.

More precisely, Williams's argument runs as follows: to believe that p is already to take p to be true, so that inquirers cannot first seek to believe things and then decide which beliefs to treat as true; inquirers interested in arriving at true beliefs must *seek* true beliefs from the first. But in a non-magical world, inquirers do not just come

by true beliefs by wanting them; they need to *operationalize* the search for truth: to develop practically workable *methods* by which to track down true beliefs. However, truth is not the kind of property that inquirers might track directly; inquirers therefore need to track some *indicator property* which recognizably indicates, if only defeasibly, that the beliefs bearing that property are true. That indicator property, Williams writes, will have to be something like '*appropriately produced in a way such that beliefs produced in that way are generally true*' (2005b, 31), where 'appropriately' is meant to rule out accidentally true beliefs, which, again from a practical point of view, inquirers cannot reasonably make it part of their *policy* to seek out, since the thing about accidentally true beliefs is that they might well have been false. Assuming that possession of that indicator property is sufficient to turn a belief into knowledge—something that KFE could accept as long as the transformation does more than just *add* something to belief—we reach the conclusion that what situated inquirers really need to seek, and what their concepts should equip them to seek, is nothing short of knowledge. Williams concludes that 'starting merely from the idea of pursuing truth in a non-magical world, and so of the truth-seeker's using methods of enquiry, we do arrive at the conclusion that the search for truth is the search for knowledge' (2005b, 31). 'In wanting the truth', the inquirer already 'wants to know' (2005b, 27). This line of argument is one way to hear Williams's remark that knowledge emerges as being prior to belief on the Craigean approach. The need for truths, when lowered into the practical context of human life where it is confronted with contingent hurdles that remain invisible from a purely abstract perspective, assumes the form of a need for knowledge.

But if our starting point is spelled out in these KFE-friendly terms, it arguably presupposes the existence of knowledge itself, and one might worry that this conflicts with the declaredly pragmatist spirit of Craig's account[8]—in particular, that it runs afoul of the pragmatist commitment to eschewing, wherever possible, the metaphysics-inviting strategy of explaining the *concept* of X in terms of the *prior existence* of X *itself*. Must pragmatists worthy of the name not put a ban on mentioning the object of the concept they seek to understand, and try to explain how the concept would come about nonetheless, driven by forces that have nothing to do with being sensitive to its object?

The simplest response would be for the pragmatist to concede that knowledge is just one of those cases where the pragmatist ambition to explain X in terms of the concept of X without drawing on X in any way reaches its limits, and that this is indeed no surprise if, as advocates of KFE suggest, knowledge is fundamental to having a mind at all (Williamson 2000, 48). Within KFE, the idea of imagining a human community in which there is no knowledge 'is probably not intelligible',

---

[8] Craig describes his method as 'pragmatic synthesis' or 'practical explication' (1990, 8, 141). He also speaks of his 'pragmatic method' (1993, 44).

because 'human beings always have perception and other automatic ways of knowing that it would be hard to imagine us...lacking for any significant length of time' (Reynolds 2017, 12).

But while it is true that *we* could not *grasp* what it would mean for a creature to have knowledge if we did not have the concept of knowledge—it would not be intelligible to us *as* knowledge—it does not follow that knowledge cannot *exist* without the concept of knowledge. Consequently, there is room for us to contemplate a situation in which agents have knowledge, but lack the concept thereof, so long as *we* possess the concept of knowledge that allows us to conceptualize the situation in this way; and we can illuminatingly do this as long as the practical contribution of the concept differs from the practical contribution of its object. This concessive response retains the compatibility of Craig's genealogy with KFE while granting that its strategy is characteristic of pragmatism only up to a point.

But there are also two less concessive responses. The first insists that Craig's genealogy does not run afoul of pragmatist principles at all: even on the KFE-friendly reading, Craig does not start by asking what knowledge *is*, or what 'knowledge' *refers* to; in good pragmatist fashion, he starts with agents and the practical needs that give point to the concept of knowledge. Moreover, as Simon Blackburn (2017) has shown, it is by no means clear what exactly the pragmatist ban on mentioning certain entities in one's explanations should amount to. What is clear is that Craig *would* run afoul of pragmatist principles if he said that we think in terms of knowledge because there is so much knowledge around—if, to put it slightly more carefully, his explanation had the following form:

There is X.

We need to be suitably sensitive to the presence of X.

Therefore, we develop the concept of X for this purpose.

But Craig explicitly rejects a strategy along these lines (1990, 3), and this is where he proves himself a pragmatist after all. It is not the object—knowledge—that *attracts* the use of the concept of knowledge. It is the agent's needs that *drive* the emergence of the concept by which the object, knowledge, can then be delineated. In other words, Craig's genealogy is subtler and more informative than the above schema because the need to be sensitive to the presence of knowledge is his *explanandum* rather than his *explanans*.[9] He derives this need from other needs instead of assuming it.

What makes model-based genealogies so useful in dealing with concepts that are fundamental to human life really only comes out once we take an even less

---

[9] I leave aside the question whether knowledge's being a social or a natural kind makes any difference to the compatibility of KFE with Craigean genealogy. I agree with Kusch (2013) and Kusch and McKenna (2020, 1065–70) that it does not.

concessive line, however. We can freely grant that if the concept of knowledge is as central to our lives as KFE claims, any state-of-nature scenario will soon covertly presuppose not just the existence of knowledge, but even the existence of the *concept* thereof: given KFE, any actual community that lacked the concept of knowledge would also have to lack a host of further concepts and practices that *depend* on the concept of knowledge, such as the practice of *telling* someone that *p* (governed by the norm that one must know that *p*), or even the concepts of *seeing* or *remembering* that *p* (which imply knowing that *p*). This is a point pressed by Elizabeth Fricker (2015) in her critique of Craig. In response, Steven Reynolds (2017) has suggested that we imagine a community deploying variants of the concepts of *seeing* or *remembering* that do not depend on the concept of knowledge—*seeing** and *remembering**, which can be used to report appearances without entailing attributions of knowledge (2017, 31–4). In principle, one could do this across the board until one had eradicated all traces of the concept of knowledge in the initial stage of one's model. But surely what matters is not that we *do* replace all affected conceptual practices with unconnected variants in our model, but that we *could* do so without losing the genealogy's explanatory force. Just because the genealogy is a model that idealizes and sometimes distorts reality, we can acknowledge that some things cannot, ultimately, be separated, and still 'fruitfully postpone considering all these things together' (2000, 153), as Williams puts it. We can instructively *separate in fiction what belongs together in reality*, thereby treating as independent what is in fact related, in order to identify what needs in particular a certain element in our conceptual arsenal answers to.

By allowing us to consider as arising separately what in reality has to arise together, pragmatic genealogy helps us organize and break down into manageable pieces the practical contribution made by concepts that are fundamental to human life. KFE claims for the concept of knowledge a status that many philosophers are willing to grant the concept of truth, namely the status of being so fundamental that we struggle to imagine language-using human societies without already covertly drawing on it. But why should the mere fact that the concept of truth is fundamental in this way bar us from using state-of-nature fictions to help us identify some of the functions that the concept of truth performs? As Miranda Fricker points out, it is precisely the 'genius of using the state-of-nature format in the arena of epistemology' that it allows one 'to tell a narrative story about X (e.g. the concept "know") even where we find it otherwise barely intelligible that there could have been a narrative development towards X', for instance because 'the idea of a progression towards X is conceptually impossible' (1998, 165).[10] The fact that a concept is so fundamental as to be involved in some of the most basic human activities only means that we would do well not to think that one simple state-of-

---

[10] Fricker notes that Williams made this point in a lecture entitled 'Truth and Truthfulness', delivered at Birkbeck in May 1997 (Fricker 1998, 165n13).

nature story will exhaust that concept's practical contribution across the entire range of human affairs. The question, then, is not whether the concept that forms the target of the genealogical narrative was covertly presupposed in some part of the state-of-nature model—if the concept is fundamental enough, it very likely was—but whether it was presupposed *in a way that renders the narrative unilluminating*. As long as the genealogy uncovers instrumental relations between the concept and certain needs that we were unaware of, it will retain its interest.

Even by pragmatist lights, therefore, Craig can illuminatingly ask what would drive a community of inquirers who need to know whether *p* to develop the concept of knowledge—even if this concept is internally related to other basic human activities, and even if this presupposes the existence of knowledge. The former is no problem because the genealogy is imaginary, and can, in virtue of this, help us identify some of the salient ways in which the concept serves our needs; and the latter is no problem because the genealogy does not presuppose the need to be sensitive to the presence of knowledge, but rather, as we shall now see, illuminatingly derives this need from other needs.

6.2 What Informants Need to Be

Let us turn to the *believing-as-the-core-of-proto-knowing* problem. If the good informant must be someone who truly believes that *p*, this puts the concept of belief at the core of the concept of proto-knowledge, and thereby also at the core of the concept of knowledge. Upon closer inspection, I want to suggest, this problem does not arise, and Craig's account even lends succour to KFE.

To see why, we need to delve further into Craig's genealogy. Given inquirers' need to find out the truth as to whether *p*, they can make some headway by relying solely on perception, reasoning, and memory—their 'on-board' resources (Craig 1990, 11). Yet the mere fact that inquirers are not all in the same place at the same time entails that there are pressures on inquirers to engage in cooperation to find things out. Someone else may have seen what I have not. Hence, inquirers need to tap into one another's stores of information: where a given inquirer's on-board resources are no help, the inquirer needs a *good informant* as to whether *p*. In characterizing good informants, Craig also tends to take his lead from the programme of conceptual analysis. He notes that they typically are individuals who truly believe that *p* and also display further properties that render them suitable as informants (1990, 12–15, 96). But what determines which characteristics are central to his account is whether they are practically relevant to the inquirer's success. 'We must never forget', he insists in an echo of Williams, 'that the inquirer's situation is a practical one' (1990, 15). From a practical point of view, good informants need to be (i) as likely to be right as to whether *p* as is necessary for the inquirer's purposes; (ii) accessible to the inquirer here and now;

(iii) intelligible to the inquirer; (iv) honest, in the most basic sense of being open with the inquirer; and (v) such that, for whatever reason, the inquirer finds their testimony convincing.

If inquirers need good informants, they also need to be able to recognize them as such. Good informants need to be identifiable as satisfying some or all of the above conditions through indicator properties: standing in the right causal relation to the state of affairs in question, for instance, or having proven reliable in the past, or being able to offer justifications. It is here that we begin to discern the roots of the vexed question of warrant, i.e. of what the 'third condition' on knowledge should be. The genealogy sheds light on this third condition 'by setting it against helpful ideas of what the point of imposing such a condition might be' (Williams 2010, 215). As Craig puts it, we hypothesize the point of something like the concept of knowledge, combine this with 'a few platitudes about the way in which human inquirers operate' (1990, 16), and generate a set of descriptive conditions characterizing the circumstances under which the concept *typically* serves that point.

Given that inquirers need to become sensitive to the presence of certain indicator properties, they need a *concept* whose application tracks these indicator properties. It is here that the prototypical form of the concept of knowledge—*proto-knowledge*—comes in: it fills the need for a concept picking out 'good informants whether *p* given the needs and capacities of the inquirer here and now'. Note that *proto-knowledge* is still markedly different from our concept of knowledge. It is strongly perspectival, i.e. indexed to the particular inquirer's needs and capacities at that time and place: a proto-knower is someone who will suit *my* needs, given *my* capacities, *here* and *now*. Moreover, the concept is what we may call *purely instrumental*: a concept is purely instrumental just in case (a) the concept is instrumental in serving the concept-user's needs; and (b) the concept is the concept *of* something instrumental in serving the concept-user's needs. The status of being a proto-knower is so closely tied to my needs that there is no conceptual room for a proto-knower who, for whatever reason, is no use to me.

We can now see that the *believing-as-the-core-of-proto-knowing* problem evaporates under analysis. This is because what the inquirer has a need *for*, on Craig's account, is not in the first instance someone who has a *true belief* as to whether *p*, but someone who *proto-knows* whether *p*, and believing that *p* is not a necessary condition on proto-knowing that *p*. It is only one among many properties that correlate well with telling the truth about *p* (1990, 13). Circumstances concurring, good informants may not need to believe what they say, and even a diffidently offered piece of information can come to be accepted by the inquirer—students might come to know the theory of evolution through their competent but creationist schoolteacher, for example (Lackey 2007). It is simply more likely that an informant who believes that *p* will come out with *p* and do so in a manner capable of persuading the inquirer. But from a practical point of view, what matters is that

the informant is likely enough to be right, accessible, intelligible, and willing to say whether p in a manner capable of persuading the inquirer.

The heart of Craig's proposal is thus the inquirer's need to identify proto-knowers. The need to identify people who believe that p comes in only derivatively, via the need to recognize proto-knowers as such. Hence, Craig's genealogy lends succour to the priority of knowledge over belief in this respect. The picture he paints is not one on which humans are first driven to develop the concept of belief and then come to delineate a privileged subset of beliefs as knowledge. Rather, the concern to identify those who proto-know whether p by deploying the concept of proto-knowledge leads the way, trailed by a derivative concern to identify typical but not necessary characteristics of proto-knowers, such as their believing that p.[11] In fact, as will become clear once we consider the second part of Craig's genealogy and the process of 'objectivization' it describes, we can even drop the qualification 'proto-', and say simply that the concern to identify knowers leads the way.

6.3 A Genealogy Showing There to Be No Room for Genealogy

This leaves the *no-room-for-genealogy* problem: if the concept of knowledge is primitive and explanatorily basic in the way Williamson suggests (2000, 185), this can seem to exclude any genealogical development towards it. Now on the face of it, there is a straightforward way of reconciling Craigean genealogy and KFE on this front. One can simply point out, as Kusch does (2009b, 90), that the concept of knowledge may be logically primitive without being genealogically primitive: the concept of knowledge may now hold a fundamental place within our conceptual scheme in a way that precludes its factorization into logically prior constituents, but this does not exclude its having developed out of genealogically prior predecessors, and it is compatible with KFE that these predecessors would have been factorizable into constituents. But of course it is open to KFE—and would certainly be in its spirit—to retort that the concept of knowledge is also *genealogically* primitive. And then KFE really would be incompatible with Craigean genealogy.

Or would it? The line I wish to press is that Craig's genealogy *itself* suggests—without becoming entangled in contradiction or forsaking its compatibility with KFE—that the concept of knowledge is genealogically primitive. Looking into the real history of the concept of knowledge, we are unlikely to find an ancestral form of it that is still far from the full-fledged thing, because nothing short of that will do. On this reading, Craig's genealogy lends succour to KFE not only by

---

[11] See Hannon (2019, 43–4) for a different development of the Craigean story according to which belief ends up being a requirement on knowledge.

suggesting that concern with knowledge is prior to concern with belief; it also buttresses and amplifies KFE's claim that the concept of knowledge is logically primitive by explaining *why* it is logically primitive and why we should expect it to be *genealogically* primitive as well.

That the affinities between Craigean genealogy and KFE should run so deep may come as a surprise even to interpreters of Craig who are sympathetic to KFE. Kusch and McKenna (2020, 1068), for example, claim that Craig's genealogy *undercuts* the claim that the concept of knowledge is primitive: by taking the failure of traditional analyses of the concept of knowledge to motivate a genealogical approach to it, they argue, Craig's genealogy undercuts a crucial motivation for KFE, for Williamson takes that same failure to motivate the claim that the concept of knowledge is primitive.

But we need not see these motivations as competing with one another; on the contrary, if the concept of knowledge is taken to be primitive, this just renders all the more pressing the question *why* this primitive notion should have been found useful, since its utility cannot then be explained in terms of the individual utility of its conceptual components. And this is precisely the question that Craig's genealogy answers: it presents the concept of knowledge as *practically basic* and explains why we should expect any human society to operate with something very like the concept of knowledge rather than with the concept of the good informant, because nothing short of the concept of knowledge will do.

There is therefore an important sense in which Craig's genealogy itself shows that there is little room for a genealogy of the concept of knowledge. There remains plenty of room for genealogy *downstream* of the emergence of the concept of knowledge, of course. But if Craig is right, there is hardly any room *upstream* of it. This facet of his account is bound to remain hidden as long as we think of genealogy as mirroring the course of history. On the interpretation of pragmatic genealogy I have given, by contrast, the method is less beholden to the course of history, and there is nothing paradoxical about using a dynamic model to show why we should expect always already to find the concept of knowledge rather than some simpler prototype of it. A pragmatic genealogy may illuminatingly *factor in sequentially* practical pressures that in reality would always already have been at work *simultaneously*. This is a capacity of the method that Craig exploits: the later part of his genealogy factors in needs that in reality do not come *later* at all, but are already at work *alongside* the needs the genealogy starts out from. This implies that earlier stages in Craig's genealogy of the concept of knowledge—the various forms of *proto-knowledge*—are still entirely free of the practical pressures of later stages in a way that no real concept could ever have been. Any real concept would face all of these pressures together, and would thus need to leapfrog Craig's earlier stages to achieve stability in the face of these pressures. This is why the form of the concept that Craig's genealogy issues in must really be the *most primitive serviceable form* of the concept of knowledge.

Nothing in Craig's account excludes our writing histories of how the concept of knowledge was elaborated downstream of that primitive serviceable form. What it suggests is that any such history will likely already *start* with something very like the generic concept of knowledge that Craig ends up with. In reality, the endpoint of Craig's genealogy must be the starting point.

By imaginatively diverging from real history as he does, Craig makes use of a power of the pragmatic genealogical method that we will find on display again in Williams's genealogy of truthfulness in Chapter 7, and it may increase the clarity and plausibility of the interpretation I propose to highlight the shared structure of these two genealogies already here. As we shall see, Williams clearly does not think that, as a matter of actual history, truthfulness first arose in purely instrumental form and only later acquired intrinsic value by coming to be regarded as a virtue. The whole point of his genealogy is rather to show that truthfulness could *only* have been stable insofar as it was valued intrinsically. Similarly, the upshot of Craig's genealogy is not that the concept of knowledge actually first emerged in its prototypical form and then developed into something else. The genealogy represents as arising sequentially what in fact has to arise together, and just as the most primitive form of truthfulness we should actually expect to find already involves intrinsic valuing, so the most primitive form of the concept of knowledge we should actually expect to find is already the 'objectivized' concept Craig reaches in the second part of his genealogy. Craig and Williams both imaginatively diverge from real history in order to convey the same type of counterfactual insight: if X *were* to arise, it *would*, because of the combination of certain practical pressures, likely be driven to develop into Y. This does not carry the implication that X in fact ever existed; on the contrary, it explains why we find Y rather than X, because it shows why, if X ever existed, it would soon have turned into Y.

This is key to understanding the second part of Craig's genealogy, which seeks to explain why we deploy the concept of knowledge rather than of proto-knowledge, and why, if the latter ever existed, it would soon have turned into the former. The concept of knowledge differs from that of proto-knowledge in various ways, most notably in that it picks out more than just good informants: someone might *know* something without being inclined or even able to pass on the information to me. Someone could be useless as a good informant, and thus fail to qualify as a proto-knower, while still qualifying as a knower according to the concept of knowledge we actually have—to take Craig's example, Luigi knows where he buried Mario's body, but he is not telling.[12]

It thus seems that, in contrast to the concept of proto-knowledge, *our* concept of knowledge does not essentially involve the notion of subservience to practical needs, which allows *us* to think what agents at the genealogical stage we are

---

[12] See Craig (1990, 17, 82).

considering cannot yet think, namely that someone could know something *without* bearing any of the marks of a good informant. This point becomes important when Craigean epistemology is criticized on the grounds that it explicates being a knower in terms of the publicly recognizable properties marking one out as a good informant. Catherine Elgin (2020, 324), for example, points out that the properties regarded in a given society as the marks of a good informant are apt to reflect the pathologies of that society; consequently, to tie the concept of knowledge too closely to these marks leads to epistemic elitism and does an epistemic injustice to those who know their stuff but lack most of the marks of a good informant—Ramanujan was a brilliant mathematician even though his lack of formal training made it difficult for his peers to identify him as such.

This worry has force if we collapse Craig's description of the concept of proto-knowledge into a description of our actual conceptual practice, as Elgin's primary interlocutor, Michael Hannon (2019), in fact tends to do. But if we insist that the genealogical dimension is introduced precisely to prevent us from mistaking a description of the concept of proto-knowledge for a description of our actual conceptual practice, we shall have more room for the thought that while our concept of knowledge is helpfully seen as rooted in the need to identify good informants, so that the concept of *proto*-knowledge initially aligns with the concept of a good informant, there are good reasons why the actual concept of knowledge would have developed beyond that stage. (Indeed, Elgin's critique presupposes that it has, since it turns on the point that someone can be a knower without displaying any of the marks of a good informant.)

Craig's genealogy explains why the concept of proto-knowledge would have turned into something more like our concept of knowledge and thereby emancipated itself from the concept of a good informant. To this end, it factors in pressures Craig takes to be acting on concept-formation more widely, in particular the pressures driving what he calls the process of *objectivization*: the more concept-users resemble social and cooperative creatures with different needs and capacities, the more there are pressures on their concepts to emancipate themselves from their practical origins as private tools answering individual concept-users' needs.[13] At the beginning of the process of objectivization, concepts are thoroughly indexed to the needs and capacities of specific agents. As objectivization proceeds, this indexation becomes weaker and the concepts less perspectival and more objectivized.

To understand what drives concept-users from purely subjectivist towards more objectivized thought, we need to understand what drives them to distinguish between the *invariable, objective properties of things* that render them useful, and the *variable, subjective needs and capacities* that incite and enable individuals to

---

[13] Here I draw on Craig (1990, 82–97) and especially on his Wittgenstein Lectures, available only in German (1993, 81–115).

use those things. It helps to consider, from a first-personal point of view, a primitive form of concept-mongering that lacks such distinctions. At this primitive stage (which Craig 'slanderously' imputes to the barnacle, as he puts it), I only have a need and am in want of something that will satisfy it, there and then. What I minimally require is the conceptual capacity to distinguish *holistically* between situations that can satisfy the need and situations that cannot. This is to wield a thoroughly subjectivized concept picking out *whatever can satisfy my present needs, given my current capacities, here and now.*

All but the simplest organisms will be driven beyond this primitive holism, however. I shall be able to exploit many more of the opportunities that the environment affords if I am capable of making more fine-grained distinctions. There are practical pressures on me to distinguish *whatever can satisfy my present needs, given my current capacities, here, now*, from whatever can do so *later*; or *somewhere else*; or *given capacities I anticipate developing* (my energy-reserves are now depleted, but I will soon recover); or given *needs I anticipate having* (I may not be hungry now, but I will soon need food again). I am thus driven to discriminate between these and many more different aspects of opportunity-affording situations and thereby become sensitive to new opportunities afforded by situations I do not currently occupy, but which I might come to occupy.

If I am a *social* and *cooperative* creature, additional advantages will come with sensitivity to opportunity-affording situations that I cannot myself occupy—those available *given someone else's needs or capacities*, for example, whom I might direct towards such opportunities in the hope of reciprocation. I will likewise be interested in the directions others can give me, and with that interest comes the interest in their operating concepts that are detached from their perspective—just as *they* have an interest in my being able to abstract from my needs and capacities when directing them. Even if I plan to free ride, I need to appreciate their points of view in order to make effective use of them. At the end of this process lie *shared* concepts that are insensitive to the differences between concept-users. They track the objective properties of things that render them suitable to certain uses, irrespective of whether anyone in particular has the need or capacity to use them. Private thinking tools have turned into public ones.

Applied to *proto-knowledge*, these dynamics driving objectivization would similarly lead to the concept's emancipation from subjective needs. A group of social and cooperative humans that started with this subjectivized concept would be driven to distinguish between the objective grounds of suitability as informants and the subjective needs and capacities to make use of them. Milestones in that process of objectivization include: (i) self-ascription, as individuals come to scrutinize their own qualifications in answer to the question 'Who knows whether *p*?'; (ii) the direction of third parties to people who might be good informants for them (note that the greater the variety of inquirers an informant has to be suitable for, the more demanding the role of the informant becomes—at the limit, an

informant would be suitable for whomever is asking, whatever their purpose); (iii) the reliance on the identification of good informants by third parties; (iv) group ventures in which individuals need not care about whether they know whether p as long as someone in the group knows.[14]

These developments would progressively weaken *proto-knowledge*'s indexation to individual perspectives and issue in something close to *knowledge*: the concept of someone who has a sufficient probability of being right about whether p for anyone's purposes, independently of the needs and capacities of any particular inquirer (if the standard for counting as a knower varied wildly from one perspective to the next, knowledge attributions would too rarely allow others to rely on them for a variety of purposes);[15] who may or may not be open, accessible, or intelligible to any particular inquirer; and who may or may not be straightfor-wardly identifiable to any particular inquirer as satisfying these criteria. Craig concludes that 'the concept of knowing…lies at the objectivized end of the process; we can explain why there is such an end, and why it should be found worth marking in language' (1990, 90–1).

Even if the concept of knowledge did not have much by way of historical predecessors, therefore, Craig's genealogy helps explain why this is so: given the practical exigencies highlighted by the dynamic model of the process of objectiv-ization, nothing short of the concept of knowledge—or at least something very like it—will do. The conviction animating the *no-room-for-genealogy* objection, that the concept of knowledge is basic, thus turns out to be underwritten by Craig's genealogy: the genealogy shows the concept to be practically basic in that any human society would find it hard to get by without it. We are social and cooperative agents who have a need for truths; hence a need for informants; hence a need to identify who proto-knows whether p; hence a need for the concept of proto-knowledge; hence a need for an objectivized form of the concept of proto-knowledge; hence a need for the concept of knowledge. Only the latter forms an apt response to the combination of needs which Craig derives from uncontroversial facts about generic social and cooperative inquirers. This makes it indeed unlikely that there should have been a gradual historical development towards the concept of knowledge.

6.4 The Art of Our Necessities Is Strange: De-Instrumentalization

What Craig's genealogy tells us goes beyond what it tells us about the concept of knowledge. It highlights far more general dynamics that help explain why concepts

---

[14] See also Fricker (2010b, 61), Kusch (2011, 9–10), and Hannon (2013, 905–6; 2019, ch. 2).
[15] See Williamson (2005, 101) and Hannon (2013, 916). For discussions of these pressures towards higher standards, see Henderson and Horgan (2015), Grimm (2015), and McGrath (2015).

are driven to shed the traces of their practical origins in individual needs. In doing so, the genealogy enriches and strengthens the methodological framework it operates in—in particular, the *instrumentalism* at the heart of this framework, i.e. the idea that our conceptual practices are illuminatingly approached as *tools* or *techniques* serving practical needs.[16]

This kind of instrumentalism about concepts faces two significant hurdles. One is that if we approach concepts as tools originating in the individual concept-user's needs, we should expect most concepts to be tailored to the individual's particular needs and capacities. Yet this is not what we find: most concepts seem to be not personalized in this way, but, as philosophers have been wont to insist, *shared*—as Frege put it, they 'confront everyone in the same way' (1893, xviii).[17]

The other hurdle is that the only form of thought that instrumentalism seems capable of explaining is instrumental thought, i.e. thought conforming to the schema 'X is a means by which to satisfy my need for Y'. But we evidently do not view the world solely from such a technical point of view. To reduce human thought to instrumental terms would be to flatten it beyond recognition. Even if our conceptual practices ever were as directly instrumental to the satisfaction of individual needs as instrumentalism represents them as being, they have since been *de-instrumentalized*, largely shedding the traces of their instrumental origins. As King Lear observed: the art of our necessities is strange.

Craig adds to the plausibility of the instrumentalist approach by showing how it can overcome both hurdles. Regarding the first, his genealogy reminds us that while it may well be true that most of our concepts are shared, this is not a conceptual truth—it is a matter of fact calling for explanation; and Craig offers such an explanation in his discussion of objectivization. It explains *why* concepts become shared by becoming progressively less tied to the specificities of particular concept-users. The publicness of concepts, though often treated as a brute fact, thus becomes intelligible as the product of practical demands arising out of two human characteristics: human life is social and cooperative; and concept-users differ in their needs and capacities. Together, these two facts generate pressures on concept-users to operate with concepts that pick out what remains constant across different concept-users' perspectives.

This is not an all-or-nothing matter, as we can conceive of partial overlaps between perspectives as well as of concepts that are only partially objectivized. Some concepts may be more objectivized while others remain tinged with sub-jectivity. The concepts at work in the natural sciences, especially physics, tend to be more objectivized than those in aesthetics, which remain firmly indexed to

---

[16] See Brandom (2011) for an overview of this instrumentalist tradition and its roots in the work of Wittgenstein and Rorty.

[17] See Glock (2009, 2010) and Prinz (2002, 14–16, 153). Even Jerry Fodor, for whom concepts are mental particulars, takes it to be a 'non-negotiable condition on a theory of concepts' that concepts be thought of as 'public; they're the sorts of things that lots of people can, and do, share' (1998, 28).

human perception. Craig offers a framework in which to think about why a concept has been objectivized to the extent that it has. Once we think of degrees of objectivization as responses to practical demands, we can reflect on which demands in particular have driven a concept to reach just the degree of objectivization it displays, and what, in a given context, the *point* is of deploying a concept that is objectivized to this degree. In legal contexts, for instance, having concepts that are strongly objectivized serves a point, since much depends on different people using the concept in precisely the same way (think of the efforts that went into objectivizing the concept of a *contract*). In the context of choosing a wine, by contrast, oenological concepts actually serve one better if they are indexed to the drinker's needs and capacities.

As for the second hurdle, Craig's genealogy actually predicts and explains the de-instrumentalization of concepts instead of being derailed by it. His genealogy suggests that there are good instrumental reasons why concepts would come to assume a form that belies their practical origins in the concerns of individual users. This is so particularly if we see Williams (2002) as continuing what Craig started. Both genealogists can be understood as describing different phases on a developmental axis of de-instrumentalization. At the axis' point of origin lies the thoroughly subjectivized and instrumentalized concept indexed to the needs and capacities of the user. At the axis' middle point lies the objectivized but still instrumentalized concept which, while public and no longer indexed to the individual user's needs, is still the concept *of* something instrumental to need-satisfaction and hence indexed to a generic subject's needs. At the axis' endpoint, finally, lies the objectivized and de-instrumentalized concept that is no longer indexed to anyone's needs, because it is no longer the concept of something instrumental to need-satisfaction at all (but, for instance, of something intrinsically valuable).

Williams's genealogical account in effect starts out where Craig leaves off. It shows why, on purely instrumental grounds, dispositions of truthfulness would need to take on more than merely instrumental value for them to possess instrumental value. While Craig's genealogy describes the dynamics driving the development from the point of origin to the middle point, thereby revealing the *instrumental value of objectivized instrumental thought*, Williams's genealogy describes the dynamics driving the development from the middle point to the endpoint, thereby revealing the *instrumental value of non-instrumental thought*. 'It is', as Williams puts it, 'not just at a technical level that we need to understand the world if we are to make a difference to it which we might recognize as an improvement' (2006i, 329–30). We have an instrumental need for non-instrumental thought.

In thinking about the instrumentality of thought, we need to disambiguate between thought that is couched *in* instrumental terms and thought that is instrumental *to* the satisfaction of needs, whatever terms it is couched in. Of

course, thinking in instrumental terms is often instrumental to need-satisfaction. But sometimes, non-instrumental thought will make the better instrument. As we saw in Chapter 3 with the example of loyalty, there are circumstances under which being bloody-minded rather than benefit-minded will reap more benefits—notably when a concept needs to throw new reasons for actions into the balance in order to make a practical difference.

Recognizing the continuity in the dynamics of de-instrumentalization described by Craig and Williams also indicates how Craig's genealogy might be pursued further to describe the development of the concept of knowledge from the middle to the endpoint of the axis of de-instrumentalization, where knowledge is conceived as possessing more than instrumental value.[18] Part of what makes informants *good* informants is that they are disposed to be accurate and sincere; but since, according to Williams (2002, 58–9), instrumental motivations are not sufficient to render these dispositions reliable in the face of the manifold temptations to be inaccurate and insincere, the dispositions to accuracy and sincerity need to come to be regarded as intrinsically valuable dispositions—as *virtues*. If Williams is right, there are instrumental reasons why the notion of goodness encapsulated in the concept of the *good* informant cannot remain simply of the kind involved in being a good knife—the goodness of a tool suited to its purpose. Rather, the concept of the good informant, and thus of proto-knowledge, needs to involve intrinsic valuation in order to sustain the institution of information pooling that gives the concept its point. This insight might be used to explain why knowledge has been deemed superior in value to true belief, and why achieving knowledge has been thought to involve the exercise of virtue.[19]

The second hurdle for instrumentalism about concepts can thus be overcome using the idea of the instrumentality of non-instrumental thought. By giving instrumentalist approaches to concepts a genealogical dimension, Craig and Williams enable us to grasp the practical origins of concepts in the needs of concept-users while also explaining why these concepts have come so far from their instrumental origins.

6.5 Assessing and Synthesizing Competing Accounts of Concepts

Finally, an important contribution of Craig's account is that he harnesses pragmatic genealogy's ability to help us assess and reconcile competing accounts of concepts. He reminds us that our concepts are, or should be, keyed to the world we

---

[18] See Kusch (2009b, 74–6).
[19] See Kvanvig (2003), Sosa (2007), Kusch (2009b), Fricker (2009), Pritchard (2012), and Hannon (2019, 63).

live in, so that understanding what concepts would serve us well in that world can give us a sense of what our concepts are likely to be, which in turn allows us to assess the plausibility and compatibility of competing accounts of them.

Craig's own genealogy informs our take on extant conceptual analyses of knowledge in several ways. First, it renders us suspicious of any account of the concept which makes it so demanding that its extension would effectively be empty. A concept that put the threshold for someone to count as knowing that *p* too high would be utterly pointless, since it would simply fail to get a grip on the world we inhabit—nothing would ever satisfy the concept. From a practical perspective, it is clear that the concept should at least *originally* permit some form of knowledge, and that the concept would have lost that original point if it now no longer did so.[20] Miranda Fricker (2008, 48) maintains that we can enlist this insight into the genealogical primacy of knowledge-permitting situations against scepticism about the possibility of knowledge. A concept of knowledge that was as demanding as the concepts involved in formulating sceptical challenges would leave inquirers seeking good informants empty-handed. The state of nature, she writes, 'explains the commonsense idea that no one can basically be a sceptic. They must be inquirer first, and sceptic second; someone committed to the practical possibility of knowledge first, and committed to undermining that possibility second' (2008, 46). As she goes on to acknowledge, however, Craig's genealogy only shows that the concept of knowledge must *originally* have been knowledge-permitting. This does not by itself exclude its subsequent development into a concept allowing us to mount sceptical challenges.[21]

Second, Craig's genealogy accounts for the variety of conditions that have been advanced to specify what must be the case for the concept of knowledge to apply. All the familiar candidates—standing in the right causal relation to the state of affairs in question, having proven reliable in the past, or being able to offer justifications—are shown to have a common practical origin in the inquirer's need to *recognize* the good informant as such.

Third, the genealogy accounts for the vexation surrounding attempts to privilege any one of these conditions at the expense of the others, because it shows that there is no practical reason for users of the concept to do so. Each of these conditions does its bit by rendering inquirers sensitive to circumstances in which informants typically are good informants, and the wider the net is cast, the greater the catch.

Fourth, the genealogy suggests that the conditions are not necessary, but typical conditions: conditions worth tracking because doing so *typically* pays off. From a

---

[20] See Hannon (2019, ch. 3) for a discussion of how Craig's genealogy informs our view of the threshold question.

[21] Fricker then goes one step further, however, claiming that 'even the sceptic cannot escape the cognitive functionality of the origin, for that scenario is still with us, at the core of what it is for us to know' (2008, 49). I discuss whether this further inference is licensed in Chapters 8 and 9.

practical point of view, what we want are concepts that are maximally cost-effective: they should strike an optimal balance between the costs and benefits of using them in the kind of world in which they are deployed (this is what it is for them to be *keyed* to the world). One notable way in which a concept can be cheap to use is by tracking properties that are readily accessible; if it tracks properties that it takes a lot of time and energy to get at, using the concept will come at a higher cost. On the benefit side of the balance sheet, a concept earns its keep by serving its point *in practice*, which usually means over a series of confrontations with real-life cases; the idea of a concept whose application conditions, when satisfied, virtually *guarantee* that the concept will serve its point sounds enticing, and evidently holds a certain fascination for philosophers; but aiming at perfection is the enemy of the good. A concept whose application conditions are such that their satisfaction virtually guarantees the presence of what the concept is supposed to help us track, but which makes it very costly, in terms of the time and effort required, to determine whether the conditions are in fact satisfied, may well prove *less* valuable in the long run than a concept that is riskier, in that the satisfaction of its application conditions only makes it *probable* that its point will be served, but that is cheaper and easier to deploy. By these measures, Craig's model suggests that tracking the wide range of properties that *typically* mark a good informant is a better bet than tracking the narrow set of properties that a good informant *necessarily* has.

Fifth, the genealogy calls into question the ambition to spell out criteria that consistently pick out knowledge even under the most freakish imaginary circumstances. By definition, we encounter freakish cases only rarely, if at all, so that they will have left less of a mark on the concept. As Craig himself notes, moreover, the unpredictability and variability of freakish circumstances means that to 'try to make a practice of detecting freakish cases' with near-infallibility would likely involve 'high costs in time and energy' (1990, 15) which are unlikely to pay off, not least because the cases are rare. Hence, he concludes that 'it must be the standard or prototypical case at which the inquirer's strategy is directed' (1990, 15). Of course, freakish circumstances may also play into inquirers' hands: 'a purpose may be achievable in unusual ways—factors which would usually frustrate it may, if other features of the situation are exceptional, do no damage, factors which are usually vital may, abnormally, be dispensible' (1990, 15). But this only reinforces the conclusion that an inquirer had better be guided by a range of typical conditions rather than a narrow set of necessary conditions, because this provides a flexibility that allows the concept pliably to accommodate varying circumstances. What philosophers should be aiming to spell out in giving an account of the concept of knowledge, therefore, are not conditions obtaining even across freakish cases, but conditions enshrining 'the features that effect realization of the purpose when things are going on as they nearly always do' (1990, 15).

More generally, it follows from the idea that we should expect and want our concepts to be keyed to the world we live in that considering their applicability to wildly counterfactual worlds is unlikely to tell us much about our concepts and their merits. Take, for example, concepts of epistemic evaluation such as those we employ in calling someone *irrational* or *unreasonable*. Let us hypothesize that these concepts perform the function of reforming bad truth-trackers by taking as proxies such indicator properties as the inability to offer a justification for one's belief. If the world happened to be full of unwittingly reliable clairvoyants, tracking the inability to offer justifications would no longer typically pick out bad truth-trackers. In such a world, our concepts of epistemic evaluation would, *ex hypothesi*, cease to function well, because they would be premised on heuristic assumptions that did not generally hold true. But given that our world does not contain unwittingly reliable clairvoyants, it is no wonder that our concepts are such that they lead us to find epistemic fault with these hypothetical characters. Our concepts are keyed to a world in which truth-tracking and reason-giving tend to go together. That they go together is a contingent fact about our world. But the fact that our concepts would not serve us well *if* that contingent fact did not obtain in no way detracts from the functionality of our concepts in *our* world.

Likewise, we should expect the functionality of concepts in our world to admit of counterexamples: cases where the conditions tracked by the concept are given but applying the concept fails to serve anyone's needs or interests. As we saw in Chapter 2, explanations in terms of typical conditions are more tolerant of exceptions than explanations in terms of necessary conditions. This fact does not reflect a lack of ambition or a willingness to settle for less than complete accounts of concepts. There is a compelling metaphilosophical rationale for it. A concept need not track necessary conditions in order to be serviceable. In fact, it would manifest a strange impoverishment in our conceptual resources if they only rendered us sensitive to perfect regularities—it would mean forsaking all opportunity to latch on to less than perfect regularities. We live in a world in which things tend to come in continua of variations on prototypes rather than in distinct series of perfect clones. In such a world, concepts tracking necessary conditions— such as the concept of a triangle in geometry—have their uses, but we should expect the vast majority of our concepts to be more attuned to protean variety than to rigid uniformity. Where there is no accessible set of properties waiting to be tracked that will *guarantee* the fulfilment of a concept's function, the obvious solution is to *build the variety right into the concept*: to track the variety of conditions under which the concept typically, though not necessarily, fulfils its function. This increases the concept's functionality by rendering it more tolerant of variation, so that it more than makes up in flexibility what it lacks in precision—just as we might prefer a mechanism that can accommodate changes in its environment as it performs a task over a mechanism whose capacities break down as soon as variation is introduced, even if the former is less clinical in its

execution than the latter. *Variation tolerance*, as one might call it, is an important desideratum on concepts in the kind of world we inhabit.

The method of pragmatic genealogy is well suited to bringing out the formative influence of such practical considerations on our concepts, and by showing how these considerations account for the internal diversity within our concepts, the method also indicates ways to *synthesize* competing accounts of concepts. Craig's project is the best example. The 'conceptual synthesis' of his title is doubly appropriate: not only does he offer an alternative to conceptual analysis by looking at how something like our concept of knowledge might be *put together* instead of analysing the concept we actually have into its constituents; he also synthesizes various accounts *of* the concept. Treating extant accounts of the concept as data, he shows that we do not need to decide between them. Each analysis latches onto some of the criteria that really do guide the concept's application. The competing accounts only make the mistake of treating these criteria as exhaustive or as necessary conditions. Craig offers a synthesis by presenting the various accounts as partial but complementary descriptions of an internally diverse conceptual practice. The debate between internalists and externalists, for example, which turns on whether people only know that *p* if they are able to articulate the justification for *p*, loses much of its point in light of Craig's account (1990, ch. 8). As Fricker puts it: 'The picture in the State of Nature is fundamentally externalist—what matters is simply that good informants come out with the truth—but then we quickly come to see the origin of internalist intuitions about knowledge' (2008, 42). We come to see that informants' ability to give justifications is typically a good indicator of their being good informants, which explains why our concept of knowledge would track this ability and hence also why the internalist analysis latches onto a real feature of our concept. Kusch and McKenna (2020, 1060–1) point out further competing theoretical intuitions concerning the concept of knowledge that Craig vindicates and reconciles: that one can know that *p* without believing that *p* (1990, 15–16); that counter-factuals play a role (1990, ch. 3); that causal relations play a role (1990, ch. 4); that the reliability of one's methods plays a role (1990, ch. 4); that standards vary with context (1990, ch. 12); and, last but not least, that all analyses of knowledge in terms of necessary and sufficient conditions admit of counter-examples (1990, ch. 6).

The promise held out by pragmatic genealogy is thus that stand-offs between competing accounts of a conceptual practice might be resolved by showing that each correctly identifies *some* property worth tracking given the concept's function, but that in light of that function, we should expect the concept to track *all* these properties, because they form the collection of conditions under which the concept typically serves its point.

In this chapter, I have examined three reasons to think that Craig's genealogy of the concept of knowledge is incompatible with KFE, and found that far from being

incompatible with it, Craig's genealogy actually supports it. Even if the concept of knowledge did not in fact have much by the way of historical predecessors, we can still give an imaginary genealogy to help explain why it predominantly exists in the form we know. I have then argued that Craig's contribution goes beyond what it tells us about the concept of knowledge, bearing on such questions as why concepts are shared, why particular concepts have been objectivized the way they have, and why concepts are internally diverse. Along the way, I have claimed that there is an underexplored continuity between Craig's and Williams's genealogies: they both describe dynamics of de-instrumentalization that combine to yield a powerful framework within which to think about concepts and their relations to our needs. But much more remains to be said about Williams's continuation of Craig's project. It is therefore to Williams that we now turn.

7

The Uses of Intrinsic Value

Bernard Williams

Bernard Williams's last book, *Truth and Truthfulness* (2002), left posterity much to puzzle over. Here is the arch-critic of utilitarianism seeking to vindicate the intrinsic value of truth in terms of its instrumental value, thereby entrenching himself squarely in the traditional territory of indirect utilitarianism. What is more, he proposes to do so using the method of genealogy, which suggests a historical approach—but he also denies that the concept of truth has a history, and prefaces his historical remarks with an avowedly fictional state-of-nature story. Unsurprisingly, this has raised questions. What separates Williams's instrumental vindication from indirect utilitarianism?[1] And how can genealogy vindicate anything, let alone something that does not have a history? These questions—and *Truth and Truthfulness* as a whole—have not received the attention they deserve. One reason for this may be that the book was in some ways untimely. Upon its appearance in 2002, it was ostensibly directed against over-enthusiastic post-modernists whose influence appeared to be anyway already on the wane: the 'subverters', 'pragmatists', and 'deniers' about truth, who cast doubt on the ideas that inquiry could be about truth rather than power, that truth was intrinsically valuable and, at the limit, that there even was such a thing as truth. But a decade and a half later, the advent of 'post-truth' politics has vindicated Williams's sense that the value of truth needed defending. And since this exploration of the point of valuing the truth is rich to a fault, it anyway rewards engagement with the puzzles it raises.

My aim in this chapter is to resolve some of these puzzles in order to develop an understanding of Williams's genealogical method that reveals it to be uniquely suited to dealing with self-effacing functionality: the phenomenon, which we encountered in Chapter 3, whereby a practice is functional only insofar as we do not engage in it for its functionality. Key to this reading will be, first, to get clear about what exactly Williams's genealogy is a genealogy *of*; and second, to read Williams as a Cambridge pragmatist—notwithstanding the fact that 'the pragmatists' is the label he gives to his opponents. Viewing Williams's genealogy as a pragmatic genealogy allows us to see it not as a piece of erudite historical stage-

---

[1] See McGinn (2003).

setting, but as a direct answer to Richard Rorty's question: why should we value the truth? While Rorty concludes that we are better off dropping the notion of truth altogether,[2] Williams's genealogy offers a contrary answer by displaying the instrumental value of valuing the truth intrinsically. The genealogy is a perspicuous derivation, from needs we have anyway, of the need for an intrinsic value of truth—and in showing the need for the value of truth to be rooted in practical exigencies, Williams proves Rorty wrong by his own lights. Using three influential challenges to Williams's genealogy as a foil, I then show in what sense an intrinsic value can be vindicated through pragmatic genealogy and how such a vindication differs from indirect utilitarianism. Finally, I highlight the parallels of Williams's genealogy with Nietzsche's 'English' genealogy of truthfulness, and argue that Williams and Nietzsche both respond to a basic predicament that truthfulness presents us with under conditions of modernity.

7.1 Truth: What Needs Defending?

What is it that Williams aims to defend against the 'deniers' by giving a genealogy? It is far from straightforward to say exactly what his genealogy is a genealogy *of*. Some commentators present it as a genealogy of *truth*, and the Italian translation of the book is even entitled *Genealogia della verità*.[3] Others urge that it is a genealogy of truthfulness *as opposed to* truth.[4] The reading I offer splits the difference: it presents *Truth and Truthfulness* as a genealogical explanation of why we might have come to value the truth intrinsically, where this means valuing the various states and activities expressive of truthfulness intrinsically.

Williams's genealogy is not focused on the *concept* of truth, because he does not think that this is what needs defending. The concept of truth, in the narrow, formal sense in which it contrasts with the *value* of truth, is a concept we cannot help but live by. As Williams repeatedly notes, the concept of truth in this sense has no history (2002, 61, 271). It is an indefinable part of a set of connected notions, such as *belief* and *assertion*, which have to be learnt together and which play a basic role in language and thought.[5] That role 'is always and everywhere the same' (2002, 61, 271). This is not an empirical claim—Williams takes it to follow from Davidson's (1990) hermeneutical insight that the concept of truth is essential to understanding something as a believing or saying that things are thus-and-so, and therefore to understanding any kind of *variation* in such believings or sayings. In recognizing something as a representation of things as being thus-and-so, we

---

[2] See especially Rorty and Engel (2007) and Rorty (1998b).
[3] See Blackburn (2002b) and Mordacci (2016).
[4] This emphasis on truthfulness as opposed to truth is found in Honderich (2003, 140), Hacking (2004, 140), Elgin (2005, 344), and Koopman (2013).
[5] See Williams (1997a, 16–19, 91n8; 2002, 45–53, 63, 84).

are always already relying on the concept of truth in this basic role. As we saw in Chapter 6, viewing the concept of truth as being fundamental in this way does not in principle bar Williams from giving a pragmatic genealogy of it.[6] But it nevertheless provides him with a reason not to give one in practice, because it entails that this fundamental concept of truth is not what needs defending.[7]

What needs defending is our *concern* with the truth about particular subject matters. For many subject matters, such as the distant past, or the intricate workings of nature, this concern has datable beginnings; the corollary is that it may also have an end. This is what motivates Williams's vindicatory project and its focus on our concern for the truth rather than our concept of truth. The contrast heralded in the title *Truth and Truthfulness* marks a distinction between a *formal* concept that we cannot help but live by, and a *social* and *historical* achievement that requires continual cultivation. We cannot give up on the concept of truth—least of all if we are determined to maintain our commitment to truthfulness, as Williams reminds the eager unmaskers and debunkers who are passionately truthful to the point of being suspicious of the concept of truth itself, because what would their passion otherwise be a passion *for*? We can, however, give up on truthfulness, in the sense of ceasing to *value* the finding and sharing of the truth in certain areas. A commitment to the value of truth of the kind expressed, for example, in 'resistance to self-deception and to comforting mythologies' is 'not given simply with the concept of truth' (2002, 18).

It is therefore our sense of the value of truth that Williams aims to defend by reverse-engineering its point for us. But what exactly is it that one values when one values the truth? Williams explains that while he is concerned throughout the book with 'the value of truth' (2002, 6), it is strictly speaking not truth itself that bears value:

> In a very strict sense, to speak of 'the value of truth' is no doubt a category mistake: truth, as a property of propositions or sentences, is not the sort of thing that can have a value. . . . The phrase 'the value of truth' should be taken as

---

[6] As a dynamic model, such a pragmatic genealogy could start off with agents represented as lacking what historical communities always already possess, and it could represent as successively emerging what in fact has to arise together. It might thereby exhibit the point of the concept of truth and the basic human needs from which it derives. As Fricker (1998, 165n13) reports, Williams took it to be a mark of the power of the genealogical method that it allows us to tell a developmental story highlighting why X should have arisen even where we find it barely intelligible that there could have been a development towards X. It would be a different enterprise again, but an equally viable one, to explain why a community that did not have a word for truth might be led to introduce one. Being able to describe things as true may function as a device of indirect reference and generalization, enabling us to endorse or repudiate claims whose content we are unable or unwilling to specify: 'I think that everything Wittgenstein said is true'; 'Beware: nothing he says will be true'. See Blackburn (2013a, 5) and M. Williams (2013, 135).

[7] Dummett's complaint that what Williams says 'comes nowhere near what is required to elucidate the concept of truth—an explanation of how we do or can acquire the concept, and an account of that in which our possession of that concept consists' (2004, 107) is thus unwarranted.

shorthand for the value of various states and activities associated with the truth. Much of the discussion will be directed to the value of what I shall call the 'virtues of truth', qualities of people that are displayed in wanting to know the truth, in finding it out, and in telling it to other people.   (2002, 6–7)

Why can truth as a property of propositions not have value? After all, to take an example Williams uses elsewhere, when someone believes against all evidence to the contrary that her child survived the crash, the truth of her belief is of immense value to that person. But for Williams, this is the value of a *particular* truth *p*, where the value of *p*'s being true just is the value of *p*, and 'this is not the value of truth, but the value of survival' (1995f, 231). Truth, as the property of a proposition, has no more value than the proposition's properties of being *well formed* or *clear*. When we say that we value clarity, what we mean is that we value *people's striving for* clarity. Similarly, when we say that we value the truth, what we mean is that we value *expressions of truthfulness*: of the desire to find out the truth, persistence in the face of obstacles to inquiry, imperviousness to wishful thinking and self-deception, care and discipline in forming one's beliefs, in preserving the truth, and in passing it on to others. The use of metalinguistic vocabulary ('true', 'well formed', 'clear') to refer to values is best understood as referring to human dealings revolving around these properties. Hence, for Williams, to value the truth is to value the various human states and activities expressive of truthfulness.[8]

This insistence that 'the value of truth' should be understood in terms of valuing the human qualities involved in being truthful, Williams notes, is an expression of his naturalistic outlook:

> People have spoken of the value of truth: is this what they had in mind? Are we right to consider only certain human attitudes toward the truth, people's dispositions to discover it and express it? My answer is yes—it is right only to consider human attitudes. Indeed, it is part of the naturalistic outlook of this inquiry that it should be seen as an exercise in human self-understanding.
>
> (Williams 2002, 60)

Williams refuses to regard as a good explanation a 'metaphysical account which represents the objects of our knowledge and their value as in themselves entirely independent of our thoughts or attitudes', or any other explanation 'that sets truth and goodness even further above us' (2002, 61). His explanation of the value of truth is an exercise in *human self-understanding* in the sense that it starts out from

---

[8] When Price (2011, 47, 139) argues that one function of the concept of truth is to encourage us to resolve disagreements in order to reap the benefits of arguing about how things are, he inflates the concept of truth to include the value of truth. By Williams's lights, Price is talking about the practical significance of truthfulness. Likewise, the 'histories of truth' offered by Foucault (1994) and Shapin (1994) are about truthfulness.

human dispositions to value the truth, and explains these in terms of their practical value *to human life*.

As a result of this naturalistic determination to understand the value of truth in terms of truthfulness, Williams's book is burdened with two potentially confusing ambiguities. One is that because the value of truth manifests itself *in* truthfulness and involves the valuing *of* truthfulness, truthfulness plays a double role as both *act* and *object* of valuation. The other is that the phrase 'the value of truthfulness' can be used in an appositional and a possessive sense. In the appositional sense, it picks out one among our values—reference to the 'value of truthfulness' is then a reference to our *valuing the truth*, manifested in our seeking, preserving, and telling it, i.e. in our *being* truthful (2002, 13). In the possessive sense, by contrast, it concerns the value that truthfulness possesses *for* us. It refers to the value *of our valuing* the truth (2002, 15).

With these clarifications in place, we can state precisely what Williams's genealogy is a genealogy of. It is a genealogical explanation of why we might have come to value the truth—where this means valuing the various states and activities expressive of truthfulness—which is given in terms of the practical value of valuing the truth. To complain, as some have done, that the book's title is 'ironically misleading' in promising a treatment of truth thus misses Williams's point.[9] It is a treatment of truth, only one that carries with it the claim that we should transpose Socratic Questions about the nature of truth into Pragmatic Questions about the point of valuing the truth: instead of gazing up at truth itself and asking what truth *is* or which *theory* of truth we should adopt, we should face the deniers with our eyes set on *human concern with* the truth, bring out how that concern relates to the rest of human psychology and to social and political concerns, and show what valuing the truth does for us. This is the task shouldered by Williams's genealogy.

7.2 The Point of Valuing the Truth Intrinsically

Williams makes use of something akin to a pragmatic genealogy tailored to a socio-historical situation in several of his later writings, though usually not under anything like that title. Yet he repeatedly finds it useful to start his reflections on liberty, punishment, and truthfulness from an idealizing description of the 'primitive basis' (2002, 21–2) of our actual situation before progressively de-idealizing by adding increasingly local practical pressures. This, he writes, can provide 'a more generic construction or plan' which 'helps us to place [what is specific to our situation] in a philosophical and historical space' (2005c, 76). We work from what

---

[9] See, e.g., Elgin (2005, 344).

he calls a very bare 'schema' of concern, a 'matrix' rendering intelligible what it is that we most basically care about in connection with a certain concept or value, and then add in 'an associated social, historical, and cultural elaboration' (2001a, 93) which this schema has received. As he also puts it, we start from a 'structural description' (1997b, 24) of a conceptual practice and its roots in certain needs, and this description then sheds light on a practice's actual history, helping us identify different instantiations of the practice in the historical record by revealing functional analogies between different phenomena. The same structural insights also indicate which developments are only weakly, and which are not at all antici-patable on the basis of the needs we started out from. Having identified develop-ments that the needs we started out from cannot by themselves account for, we are led to look, as Williams does, for the historical moments around which further needs or pressures came into play that explain how we got from the basic schema of concern via a particular historical elaboration of it to the practice we know. Together, the basic schema of concern and its historical elaboration yield the 'contour' of a concept or value 'in a given society or historical situation' (2001a, 94).

Given that the contours of Williams's own genealogy of truthfulness are notoriously elusive, it is worth reconstructing it in some detail. (For ease of exposition, I shall work with the terminology Williams uses and not rephrase everything in 'knowledge-first' epistemology-friendly terms again.) Where appro-priate, I shall draw on Williams's other book on the value of truth—unavailable in English—which is unambiguously titled *Der Wert der Wahrheit* (1997a). In that book, Williams starts his genealogy at the beginning, namely with what he takes to be the most primitive and generic form of truthfulness. (In *Truth and Truthfulness*, by contrast, he starts later in the genealogical story and then cuts back to this explanatorily prior stage in the fourth chapter.) This most primitive form of truthfulness is the constitutive norm of belief and assertion that we need already insofar as we need to represent the world as being a certain way at all.

(1) *The Need for Thin Truthfulness as a Constitutive Norm of Belief and Assertion*: The starting point of Williams's genealogy is an abstractly characterized situation of language acquisition. Its purpose is to highlight the internal relations between truth, assertion, and belief, and to situate truthfulness in relation to them. Using Williams's own terms (1997a, 18; 2002, 94), we may call it the *situation of primitive openness and immediacy*: a speaker utters a sentence *S* which (a) describes the immediate environment, (b) is true, and (c) can readily be seen by the hearer to be true—it is not just true, but 'plainly true' (2002, 45). The situation is one of *immediacy* insofar as the sentence describes what can be seen by both speaker and hearer to be the case now and around here (rather than a state of affairs inaccessible to the hearer at the time and place of utterance). It is one of *openness* insofar as the speaker uninhibitedly speaks his or her mind and the truth lies open to view. Even this simplified situation already draws on the triad of *belief*, *assertion*, and *fact*. Following Davidson (1990), Williams holds that these three

elements must originally be lined up for understanding to get off the ground (what is believed and asserted is what is the case). But as understanding progresses, the possibility of misalignment between these three elements comes into view—there can be 'gaps' between 'the assertion of the speaker, the truth conditions of the assertion, and the belief of the speaker' (1997a, 17). There is the possibility of discrepancy between belief and fact, i.e. between how the speaker believes things to be and how they in fact are. And there is the possibility of discrepancy between belief and assertion, i.e. between how the speaker believes things to be and how the speaker's assertion represents them as being—the speaker might not say what he or she takes to be the case.

It is with this possibility of misalignment between belief, assertion, and fact that conceptual room opens up for the most primitive form of truthfulness: the dispositions involved in aligning belief with fact and assertion with belief. This suggests that truthfulness falls into two bundles of dispositions: what Williams, capitalizing the terms to mark their technical nature, calls the dispositions of 'Accuracy', which dispose one to believe things to be as they in fact are, and the dispositions of 'Sincerity', which dispose one to assert things to be as one believes them to be (1997a, 18).[10]

This first stage shows that truth stands in an *internal connection* to belief and assertion (1997a, 16–18; 2002, 45–53): both beliefs and assertions, in their different ways, *aim at truth*, and falsehood is an objection to them. This captures an important but thin sense in which truth has value: it functions as a constitutive norm of the 'belief-assertion-communication system' (2002, 84). We cannot make sense of beliefs and assertions without seeing them as subject to the norm that they should, in the basic case, be true. Insofar as we represent the world at all, we need to subject ourselves to the norm that beliefs and assertions aim at the truth; and insofar as to do this is to value the truth—to manifest a form of truthfulness—we all need to value the truth in this thin sense.

However, Williams also insists that this insight into the need for a thin form of truthfulness will not take us very far. In particular, it 'takes us nowhere at all' (2002, 84) when the question arises *whether we should continue to work that communication system in relation to certain subject matters*, such as historical narrative or political interpretation. The need for thin truthfulness does nothing to answer the deniers' worries about truth in these connections (2002, 5). Nor does it tell one whether, in a given case, one should be truthful or not. It may well be true that 'no-one can learn or speak a language unless a large class of statements in that language is recognized to be true' (2002, 5). But this does not stop people from lying; on the contrary, it is what enables lying in the first place, since lies depend

---

[10] The explanatory primacy of dispositions here mirrors Williams's approach to ethical thought. He views dispositions as 'basic' to it and maintains that 'the replication of ethical life lies in the replication of dispositions' (2006h, 75).

on there being robust expectations that utterances will be true. We know that if we lie, 'the heavens will not fall; if the heavens were going to fall, they would have fallen already' (2002, 85). To answer questions about whether we should continue to 'work the system' in a given case or area of discourse, a further, thicker sense of valuing the truth is required.

(2) *The Need for Information*: Williams works towards this thicker form of truthfulness by building on Craig's story. Whatever else we take humans to need, it will—as a matter of structural necessity—be the case that each individual has a need for 'information about the environment, its risks and opportunities' (2002, 58). This already indicates a further, if limited, respect in which individuals need to possess the quality of Accuracy: they need to be Accurate in acquiring information about their environment through their primary sources of information, i.e. their own senses and powers of reasoning.

(3) *The Need for Informants*: Williams then argues that each individual also has a strong interest in exploiting *secondary sources of information* by tapping into others' stock of information. This follows already from the idea of a *purely positional advantage* (1997a, 19; 2002, 42). Since we are not all at the same place at the same time, but continually take up more or less different points of view, we acquire different information over time. Someone sitting atop a tree may see approaching predators that I cannot see, while I may know what happened here yesterday when others do not. Already in virtue of the fact that we occupy different points of view, therefore, we sometimes need information that others are better placed to acquire than we are, and we therefore have a need for informants.

(4) *The Need for a Division of Epistemic Labour*: From (2) and (3), Williams derives the *collective need* for a practice of epistemic division of labour whereby information is shared, and each group member functions as a *contributor* to a joint pool of information.

(5) *The Need for Good Contributors*: Williams then parts company with Craig by focusing not on the demands on the inquirer, but on those on the *contributor* to the pool of information. The need to pool information in turn generates a need to cultivate the qualities that make *good* contributors to the pool of information. Most basically, these must be the qualities involved in getting one's beliefs right— for otherwise the pool is vitiated by misinformation—and the qualities involved in passing on one's beliefs to others—for otherwise one's information fails to enter the pool. It is here that the need to cultivate *Accuracy* and *Sincerity* in some thicker form comes into view.

(5a) *What Accuracy Needs to Be*: Once information is pooled, each individual has an interest in others having the quality of Accuracy as well. What can the practical demands on good contributors tell us about the contours of Accuracy? Clearly, what we most basically need from contributors to the pool is that they contribute truths to it. Displaying Accuracy must therefore involve taking care in

forming one's beliefs so that they may be true beliefs.[11] This means that contributors need to use methods of inquiry that are *truth-acquiring*—augury, or any method that randomizes over the selection of beliefs, such as guessing or pulling answers out of a hat, is not truth-acquiring (2002, 132). Which qualities, Williams asks, do we need to cultivate in people to be able to rely on them to use truth-acquiring methods?

For one thing, they should *want to find out the truth* on the question at issue. This, Williams argues (2002, 133), is equivalent to wanting to get into the following condition:

if p, to believe that p, and if not p, to believe that not p.

This formula brings out that pursuing the truth is a matter of controlling one's belief-formation. Because there may be a time-lag between the situation in which a belief is formed and the situation in which it is fed into the pool, moreover, there is also a need for the *preservation* of truth. This is where the qualities involved in remembering accurately come in, such as resistance to forgetfulness and memory distortions. Accuracy, in an apt phrase that Williams adopts from Isaiah Berlin, is the set of dispositions that secure and sustain our *sense of reality*.

This sense of reality then needs to be defended against *external* and *internal* sources of error. External sources of error include the fact that the truth is sometimes hard to find, or that finding it may require persistence in the face of obstacles. It is not always clear in advance what these obstacles are, and how much effort will be required to find out the truth on a given matter. Avenues of inquiry may remain unexplored without it being clear *what* is left unexplored. This makes it difficult to decide when *enough* pains have been taken to find out the truth. It is here that internal sources of error come into play: it may simply be more comfortable to stop there, or one may be tempted to do so because the belief arrived at is agreeable. An Accurate inquirer needs to possess the discipline to resist these temptations.

Another internal source of error that our sense of reality needs to be defended against is temptation to self-deception, i.e. the tendency to let wishes distort one's judgement. Our misapprehension of reality is then primarily due to the distorting influence of our own motivational set: beliefs become motivated by the wish, hope,

---

[11] Is it so clear that belief-formation must aim at *true* beliefs? Can we not maintain with Rorty (1998a, 19) that since we cannot tell the difference between our believing with sufficient justification (or reaching reasonable agreement) that p and its being true that p, we had better admit that our aim is just to attain the psychological or social state in question? Williams thinks not, on the grounds that descriptions such as 'believing with sufficient justification' or 'reaching reasonable agreement' *already* call upon the notion of truth: 'A justified belief is one that is arrived at by a method, or supported by considerations, that favour it, not simply by making it more appealing or whatever, but in the specific sense of giving reason to think that it is true' (2002, 129). See Dummett (2004, ch. 6) for a nuanced discussion of Williams's response.

or fear that *p*. But beliefs are answerable to the world, and cannot properly be dependent on our will in this way. When, unbeknownst to us, they are the product of wish-fulfilment, they flaunt this requirement and make us lose our hold on reality. Of course, when they concern states of affairs that are themselves dependent on our will, beliefs do properly depend on our will: whether I go for a walk today depends on my will, and my belief about the matter quite properly follows wherever my will leads. But even in these cases there is such a thing as losing one's sense of reality: there are aspects of the world that are subject to my will and aspects that are not, and my beliefs only properly follow my will when the willed states of affairs are ones I know I can affect.[12]

Finally, an internal source of error on which Williams is particularly illuminating is wishful thinking, which he characterizes as the tendency to let wishes become beliefs. The contents of the mind, Williams insists, are not exhausted by beliefs and desires:

> it is a misunderstanding of one-person practical reasoning—one encouraged, certainly, by political models of deliberation—to think of it in terms of a set of formed and committed desires adjudicated in the light of formed and committed beliefs. Rather, the process of arriving at a practical conclusion typically involves a shifting and indeterminate set of wishes, hopes, and fears, in addition to the more clearly defined architecture of desire and belief. If we think only in terms of desire and belief, we may well overlook subtle problems within the economy of desire itself, notably the fact that to distinguish between a desire and a mere wish is an achievement, and, to a significant degree, a cognitive achievement. For that reason, and more generally because of the discipline that is involved in maintaining the barriers between the route to desire and the route to belief, we can recognize that the virtues we need in considering what to do coincide at deep levels with the virtues that we need in inquiring into anything, the virtues of truth.   (Williams 2002, 198)

Wishful thinking involves not so much the misapprehension of reality as the illicit intrusion of fantasy into our stock of beliefs and desires. On Williams's picture of the mind, our beliefs and desires are formed as much out of the maelstrom of inchoate and passing wishes as out of impressions of the external world. Yet the beliefs and desires generated from wishes only maintain our sense of reality when they stand in the right relation to the world and to our will, i.e. when the beliefs are independent of our will and the desires are for things we can affect.

Part of what Accuracy needs to be, then, is a set of strategies and attitudes enabling the inquirer to resist the various expressions of what Williams, with a

---

[12]   A complication is that what is possible for me is also a function of my desires. See Williams (2002, 195–6).

nod to Freud, terms the 'pleasure principle' (2002, 125): Accurate inquirers need to want to believe the truth, to preserve it, to apply truth-acquiring methods, and to resist everything from laziness and the desire to believe the agreeable to the temptation to let their wishes distort their judgements (self-deception) or become beliefs (wishful thinking).

(5b) *What Sincerity Needs to Be*: Besides generating a need for Accuracy, information pooling also generates opportunities and motives for deception and lying. Individuals might gain from withholding information, or from sharing information that is misleading or false—hence the strong collective need to cultivate openness towards inquirers, willingness to contribute to the pool of information, and resistance to the temptations to lie, mislead, or deceive. These are the qualities that make up Sincerity.

How much does Sincerity need to involve? Clearly, it is not enough for informants to possess the disposition not to assert what they do not believe— they also need to be disposed to come out with what they believe, and to do so in a helpful way. Just as Craig argues that the concept of knowledge gets caught up in the dynamics driving objectivization, Williams argues that Sincerity gets caught up in the dynamics driving *cooperation*, and in particular the dispositions of *trustworthiness*. We saw in stage (1) of our reconstruction of Williams's genealogy that a minimum of trust in the fact that one will say what one takes to be true is a structural requirement on communication. Analogously, there is a structural requirement on cooperation that one can be trusted to behave as one leads others to expect one will behave: trust is required for cooperation in all its forms, including the division of epistemic labour introduced in (3). By *trust*, Williams means the willingness of one party to rely on another party to act in certain ways. By *trustworthiness*, he means the disposition to do what one is expected to do *because* one is expected to do it (2002, 88).[13] The two are connected, because whenever cooperation requires reliance on others, each party requires some form of assurance that the other party will not default. This gives rise to the problem modelled in game theory as an 'Assurance Game' or 'Stag Hunt': necessarily cooperative enterprises such as stag hunting require some sort of assurance that the participants can be trusted to do their bit by remaining at their post even when that means forsaking opportunities to pursue passing hares that they could catch by themselves.[14] The need for trustworthiness thus derives from a more general need for cooperation.

These requirements on trustworthiness shape what Sincerity needs to be. Sincerity is itself a form of trustworthiness, namely *trustworthiness in speech*.

---

[13] See Williams (1988) for his account of trust. See Faulkner (2007) for an elaboration of Williams's genealogy that focuses on the emergence of trust.

[14] The stag hunt imagery goes back to Rousseau (1977, III). For discussion of the Stag Hunt problems at the origin of a variety of social structures, see Skyrms (2004).

When communication takes place under circumstances of trust, hearers will be willing to rely on what speakers said. This presupposes that the speakers are trustworthy. But trustworthiness involves 'more than the avoidance of lying' (2002, 97).[15] As the story of Saint Athanasius illustrates, there are plenty of ways to mislead and deceive without lying. In the words of Peter Geach, the Saint was 'rowing on a river when the persecutors came rowing in the opposite direction: "Where is the traitor Athanasius?" "Not far away," the Saint gaily replied, and rowed past them unsuspected' (1977, 114).

Trustworthiness in speech demands more than just the eschewal of lying because 'in relying on what someone said, one inevitably relies on more than what he *said*' (Williams 2002, 100). At stage (1), we focused on the relation between the assertions and beliefs of the speaker. Complications are introduced once we bring in the *production of beliefs in the hearer*. A hearer gathers more from an assertion than just the content of the assertion itself. Speakers have countless beliefs, and countless ways of expressing them. The fact that a speaker said what she said rather than something else—and said it in one way rather than another—itself conveys information via what Grice called 'conversational implicatures' (1989).[16] Whenever a speaker expresses one belief, the hearer acquires many (Williams 2002, 100).

Consequently, Sincerity needs to be extended beyond the avoidance of lying if it is to constitute a form of trustworthiness.[17] It must grow into something more than a disposition 'to make sure that any assertion one makes expresses a genuine belief' (Williams 2002, 97); it needs to become a disposition to express *those* among one's beliefs, and in as *helpful* and *non-misleading* a form as one can be expected to do in a given situation *because* one is expected to do so.

From these reflections on what Sincerity needs to be, Williams concludes that it is a mistake to *fetishize* assertion, as he takes the Christian tradition from Aquinas to Kant to have done (2002, 95). Christian doctrine 'makes the assertion into a fetish by lifting it out of the context in which it plays its part and projecting onto it in isolation all the force of the demand for truthfulness' (2002, 107). The pragmatic reconstruction of Sincerity, by contrast, suggests that the demand for truthfulness must extend to all forms of deceit—'something is wrong if one thinks that it is more honourable to find some weasel words than to tell a lie' (2002, 107), for if 'lying is inherently an abuse of assertion, then so is deliberately exploiting the

---

[15] Where a lie is 'an assertion, the content of which the speaker believes to be false, which is made with the intention to deceive the hearer with regard to that content' (Williams 2002, 96).

[16] Williams gives two of Grice's illustrations: 'someone who says, "I went into a room yesterday and spoke to a woman…", would normally be taken to imply that this woman was not, for instance, his wife, and that the room was not in their house; while if he said, "I broke a thumb yesterday", he would, on the contrary, be taken to mean that it was his thumb (it takes on an unexpected ring if it is heard as a wrestler's boast)' (2002, 99).

[17] In *Der Wert der Wahrheit*, Williams speaks of 'extended Sincerity' (1997a, 24).

way in which one's hearer can be expected to understand one's choice of assertion' (2002, 107).

(6) *The Need for Intrinsic Values*: The story so far leaves truthfulness with certain limitations. First, its value is purely *instrumental*: it is 'entirely explained in terms of other goods, and in particular the value of getting what one wants, avoiding danger, mastering the environment, and so on' (Williams 2002, 58); second, this instrumental value is largely the value it has *for others*—it is in the interest of each individual that others reliably display the qualities of Accuracy and Sincerity, but there is, at this point, little to indicate that it is in the interest of each individual to consistently do the same—they are better off keeping information to themselves, or sparing themselves the effort of acquiring it when it does not serve their own ends. What is to stop them from being deceptive or lazy when they can reap some extra benefits that way while still profiting from the communal practice of truthfulness? Together, these two facts engender what, in game-theoretical terms, is called a *free rider problem*: 'each participant wants there to be a practice in which most of the others take part, without, if he can get away with it, taking part in it himself' (Williams 2002, 58). As a result, nobody ends up being reliably truthful.

On a purely instrumental understanding, therefore, truthfulness is unstable because its value largely lies in the value it has for other people, which renders it overly vulnerable to free riders. And this in turn means that truthfulness fails to provide a solution to the Assurance Game: in relying on what an informant tells them, others need assurance that the informant is sufficiently Accurate and Sincere; but given the temptation to free ride that purely instrumentally understood dispositions of Accuracy and Sincerity are vulnerable to, such assurance is lacking.

These problems are particularly pronounced where Sincerity is concerned, but to a lesser extent they also affect Accuracy. While in some cases the informant's own interests may line up with the interests of those who rely on him or her, there comes a point where Accuracy's instrumental value to the informant is outweighed by other considerations (the manifold temptations of the 'pleasure principle'): insofar as inquirers in the state of nature are modelled as *minimally reflective agents* that are able to weigh the expected value of prospective information against the costs of acquiring that information, they realize that collecting information has a price—it involves an *investigative investment*. The various obstacles to such an investigation may then provide reasons not to exercise Accuracy in some cases. And in these cases, the value of Accuracy will, like the value of Sincerity, lie largely in its value to others. Consequently, 'no society can get by . . . with a purely instrumental conception' (Williams 2002, 59) of Accuracy and Sincerity.

Williams thus reaches the need for Accuracy and Sincerity to turn from mere dispositions into *virtues*: dispositions valued *for their own sake*. For investigative

investments to be reliably made even beyond the point at which Accuracy serves the individual, truth needs to carry its own weight on the balance of reasons. Accuracy needs to be elaborated into the striving for truth 'for its own sake' (2002, 126). It needs to be de-instrumentalized and elaborated into what Williams calls the 'passion for *getting it right*' (2002, 126).[18] Similarly, for contributors to be Sincere even when it involves forfeiting opportunities to deceive for personal gain, 'there should be some dispositions to think that telling the truth (to the right people, on the right occasions) is in itself a good thing' (Williams 2014g, 408). Only if Accuracy and Sincerity outgrow a merely instrumental conception of them, so that virtue becomes its own reward, can they be stable enough to sustain a cooperative practice of information pooling—only in communities in which people value truthfulness for its own sake can truthfulness fulfil its function.[19]

More precisely, Accuracy and Sincerity need to fulfil the following two conditions:

(IV)   *The Intrinsic Value Condition*: X is seen by group members as an intrinsic value.

(K)   *The Knowledge Condition*: It is generally known within the group that (IV) is the case.

Fulfilling these two conditions solves the Assurance problem: inquirers will have some assurance that informants are Accurate and Sincere once informants have *internalized* the dispositions of Accuracy and Sincerity and it has become *common knowledge* that they have internalized them. (IV) and (K) together meet these conditions. They also solve the free rider problem: if people come to think that it is a good thing to act as an Accurate and Sincere person acts *just because* that is the kind of action it is, then they will have a reason to be Accurate and Sincere that is independent of the reasons that there are for others to want them to be Accurate and Sincere.

This crucial step in the genealogy shows, then, that it is only *insofar* as truthfulness is seen as possessing *intrinsic* value that it also possesses *instrumental* value. It needs to outrun its function in order to be functional. This requirement itself flows from functional considerations, which is why we are dealing here with the structure we introduced in Chapter 3 as self-effacing functionality.

But how do the dispositions of truthfulness come to acquire intrinsic value? Williams gestures at some of the social and psychological mechanisms that

---

[18]   *Pace* Finnis, who writes that 'Williams never recognizes' that Accuracy 'is necessarily a secondary element in the disposition or desire to have true beliefs. More primary is curiosity, the desire to learn, to find out. And most primary is the insight that knowledge is a good, pursuit-worthy for its own sake' (2011, 97).

[19]   Kolodny (2005, 543) also discusses the problem. Like Williams, he concludes that what is needed is an intrinsic reason to follow the rule on each and every occasion (2005, 544–5).

shoulder this task. He notes that 'people may be discouraged or encouraged, sanctioned, shamed, or rewarded' (2002, 44) into valuing truthfulness intrinsically; that the 'motivations of honour and shame play an important part' (2002, 120) and that a driving force of the process is the 'fear . . . of disgrace in one's own eyes, and in the eyes of people whom one respects and who one hopes will respect oneself' (2002, 116). The structure 'of mutual respect and the capacity for shame in the face of oneself and others', he writes, 'is a traditional, indeed archaic, ethical resource, but it is still very necessary' (2002, 121). Williams's explanation thus draws on more basic psychological and social resources that we have reason to draw on anyway elsewhere, like the emotions of fear and shame and the human susceptibility to internalize norms as a result of systematic sanction.[20]

(7) *The Need to Make Sense of Intrinsic Values*: For (IV) to hold—for something to be an intrinsic value—it is not enough to recognize a necessity to *treat* it as an intrinsic value. In addition, people need to be able to *make sense* of it *as* an intrinsic value 'from the inside' (2002, 91), as Williams puts it. That is to say, there must be more to the intrinsic good in question than just the bare necessity of its being seen as an intrinsic good. It also needs to be conceptually articulated—to possess content in a way that allows one to say what is valuable about it and how it relates to other things that one values.[21] In ancient Greece, truthfulness was made sense of in terms of ideas of honour and shame; at other times, truthfulness was connected to freedom, absence of manipulation, equality, and self-respect (Williams 1997a, 26; 2002, 115). Moreover, as Williams influentially highlighted, an intrinsic good needs to be able to engage people's emotions.[22] We cannot make sense of something *as* an intrinsic good if it remains utterly unrelated to anything else that we value and there is nothing to be said about what makes it valuable. Isaiah Berlin offers an illustration of just this point:

> If I find men who worship trees, not because they are symbols of fertility or because they are divine, with a mysterious life and powers of their own, or because this grove is sacred to Athena—but only because they are made of wood; and if when I ask them why they worship wood they say 'Because it is wood' and give no other answer; then I do not know what they mean.
>
> (Berlin 1997, 10)

---

[20] See Kusch (2009b, 77–8) for further discussion.

[21] As Kusch (2009b) points out, this indicates a contextualist conception of intrinsic value that differs markedly from, for instance, G. E. Moore's, who tries to identify intrinsic goods by asking whether they are 'such that, if they existed by themselves, in absolute isolation, we should yet judge their existence to be good' (Moore 1993, 236–7).

[22] Williams (1973d). See Goldie (2009) regarding the connection Williams draws between values and emotions and Kusch (2009b) regarding Williams's view of values as coming in socially shared webs.

And just as we cannot make sense of an intrinsic good in other societies unless we are given some more material in terms of which to make sense of it, we cannot make sense of something as an intrinsic good merely on the basis of the thought that it would be instrumentally useful for us to value it intrinsically. This instrumental-minded justification does nothing to flesh out the intrinsic value it recommends. If we do not independently possess some material in terms of which to make sense of it as an intrinsic value, the resulting structure will be unstable under reflection. As Williams puts it, the 'lack of fit between the spirit of justification and the spirit being justified is so radical that, if the construction is exposed to reflection at all, it is bound to unravel' (2002, 91). This is why Williams concludes that the recommendation of non-utilitarian rules on strictly utilitarian grounds must be unstable under reflection (Smart and Williams 1973; Williams 2011). It is also why he is unconvinced by David Gauthier's (1986) suggestion that players in the Prisoner's Dilemma might reason their way out of the Dilemma by recognizing their instrumental interest in coming to value justice intrinsically (2002, 91). If this is all they have to go on, their attempts to value justice intrinsically are bound to deteriorate into mere pretence. It is as if one were given an instrumental justification for the attitude of valuing wood intrinsically, but, like Berlin's wood worshippers, remained unable to say anything when pressed to explicate wherein the intrinsic value of wood consisted.

On Williams's account, then, an agent can make sense of something as having intrinsic value only if the agent possesses 'some materials in terms of which he can understand this value in relation to other values that he holds' (2002, 92). He perhaps too succinctly sums this up by saying that he takes the following two conditions to be sufficient for X to *be* an intrinsic good: first, it is necessary (or nearly necessary) for basic human purposes and needs that human beings should treat X as an intrinsic good; and second, they can coherently treat X as an intrinsic good (2002, 92; 2006f). In light of the above, we can unpack this by saying that (IV) divides into the following two conditions:

(PN)   *The Practical Necessity Condition*: It is necessary (or nearly necessary) for basic humanpurposes and needs that human beings should treat X as an intrinsic good, which is to say that it is a practical necessity for them to treat X as intrinsically valuable.

(SM)   *The Sense-Making Condition*: They can coherently treat X as an intrinsic good, which is to say that they can make sense of X as an intrinsic value from the inside. This condition in turn divides into:

  (i) *The Conceptual Articulation Condition*: They have the conceptual resources necessary for them to relate X to other things that they value, which is to say that X must have an inner structure, i.e. it must be conceptually articulated

in a way that allows it to be related it to other goods (as instantiating, expressing, or furthering them in some respect).

(ii) *The Emotional Engagement Condition*: They have the emotional capacities and dispositions necessary for X to engage their emotions.

These conditions are probably not best read as necessary conditions for being an intrinsic good—Williams is thus not committed to the claim that *all* intrinsic goods are goods that humans *need* to treat as intrinsic goods. But he takes them to be sufficient conditions, which, where they hold, can help explain why something is an intrinsic good.

(8) *Three Types of Historical Elaboration*: While the state of nature can reveal that there is a need for truthfulness to be valued intrinsically, and hence also for the conceptual and emotional resources this requires, it cannot show how these generic requirements have actually been satisfied now and around here. Nor can it help us understand the 'enormous degree' to which truthfulness was 'changed, transformed, differently embodied, extended and so on by historical experience' (Williams 2007, 132). A perspicuous representation of what truthfulness does for us needs to be enriched with historical, social, and cultural information that is invisible a priori. We need to incorporate increasingly local needs into our model of truthfulness' development in order to account for the elaborations that our instantiations of the virtues of Accuracy and Sincerity have undergone. Generic needs cannot account for the extension of truthfulness to the distant past around the time of Thucydides (Williams 2002, ch. 7), or for the fashioning of Sincerity into the ideal of personal authenticity in the course of the Romantic era (2002, ch. 8). Moreover, while the model shows that truthfulness needs to be valued intrinsically, it cannot help us specify the material in terms of which *we* came to make sense of truthfulness. Hence, 'philosophy, in order to do its business, must move into history' (2002, 173).

This is not to say that Williams's genealogy suddenly becomes 'history, correctly practised'. To take the first half of Williams's book to be a state-of-nature fiction and the second part to be regular historiography is to invite concerns about how such disparate genres of genealogy could coherently be stitched together into one story—concerns voiced by Rorty, for example, who remarks that he 'had trouble seeing the continuity between the first half and the second half of Williams's book; the connections between the more philosophical part and the more historical part are not perspicuous' (2002). That same worry is echoed by Koopman (2013, 68). And indeed, as long as the only two genres of genealogy we recognize are Hobbesian state-of-nature fiction and Foucauldian historiography, Williams's genealogy is bound to seem incongruously chimerical.

But once pragmatic genealogy is available as a third option, it becomes clear that there is nothing incongruous about starting in a highly idealized state of nature and gradually de-idealizing that model in the direction of our actual

historical situation. Williams does not claim to offer anything like the historian's craft in the second half of his book; he wants only to 'mention a few of the historical divergences' and to 'trace some features of the structure that give rise to the variations' (2002, 95) of the prototype of truthfulness outlined in the state of nature. To model the elaborations that truthfulness has undergone beyond its most generic form, the genealogical model needs to factor in practical pressures that are socio-historically local. But even in the second half of the book, Williams's genealogy remains an exercise in philosophical model-building throughout. While informed by history, that later part of the genealogy continues to respond to the primarily philosophical concerns that initially prompted the examination of the state-of-nature model. It is just that the genealogy needs to draw on history in order to model three types of elaboration: (a) the *specification* of the materials in terms of which the abstract framework outlined in the state of nature is filled in; (b) the *extension* of the prototype; and (c) the *autonomous elaborations* of the prototype that are contingent relative to the initial dynamics outlined in the state of nature. Let us take a closer look at each type of elaboration.

(a) *Specification*: History is required to specify the materials in terms of which the abstract framework outlined in the state-of-nature model is filled in. The state-of-nature model itself reveals motivational gaps in the prototype of truthfulness it sets out: like many other functional concepts or dispositions, truthfulness must come to be understood non-functionally by those who exhibit it. While this can be achieved in any number of ways, history is required to understand how this requirement was met in any given case. That this elaboration takes place is both *anticipated* and *necessitated* by the state-of-nature model, much as an equation may leave it open which value a variable takes while making it clear that it cannot be null.

Williams points to ideas of *nobility*, *honour*, and *shame* as traditional ways of providing the value of truthfulness with an internal structure (a case in point is the young noble Neoptolemus who, in Sophocles' *Philoctetes*, recoils from deceiving for gain because it is 'dishonourable').[23] In hierarchical or aristocratic societies, or 'in association with a very highly cultivated aesthetic' (Williams mentions Oscar Wilde), truthfulness is connected to 'motivations of a self-sufficient nobility'—the person 'who is for others, on this line of thought, is no-one in particular' (2002, 116). In Enlightenment thought as exemplified by Kant, by contrast, we find a poignant articulation of a connection between deceit and manipulation: the use of others 'merely as means' (2002, 119). The value of truthfulness is then primarily made sense of in terms of 'individual freedom and the avoidance of manipulation' (2002, 122).

---

[23] See Williams (1997a, 26; 2002, 115).

In the case of Accuracy, too, ideas of 'conscience, honour, or self-respect' (2002, 126) have been drawn on to lend it reflective stability, and the story of how the idea of Accuracy came to be filled in is largely the story of 'the cultural and eventually industrial sophistication of this idea into what is now called "science"' (2002, 141).[24] The disposition to Accuracy is fashioned into the scientific striving for 'a representation of nature which abstracts to the greatest possible degree from the perceptual and other peculiarities of human beings' (2002, 143).[25] With a nod to Max Weber's *Wissenschaft als Beruf* (1922), Williams outlines how 'the search for truth becomes...an intrinsic good', which in turn enables 'the desire of a scientist to discover and hold on to reality' to 'stand against such forces as political corruption and terror' (2002, 141). Scientific inquiries come to be seen as expressions of freedom. In striving to filter out the contribution of humanity to experience and confronting the inquirer with something that is not 'a rival will' (2002, 146), they offer a kind of 'liberation from humanity' (2002, 143). These are some of the ways in which people have made sense of truthfulness from the inside.

(b) *Extension*: Second, history is required to understand how the prototype might have been extended through developments that are anticipated but not necessitated by the state-of-nature model. After all, the needs sketched so far only demand a restricted form of truthfulness (2002, 40). They suggest that one would need to tell the truth to members of one's epistemic community, for instance, but not to everyone. And this is indeed what we find in the world of Homer, where friends deserve the truth while enemies do not (2002, 121–2). In the world of Rousseau and Kant, by contrast, where truthfulness becomes connected to ideas of moral equality, this leads to the circle of those deserving of the truth being extended (1997a, 27; 2002, 117). At the limit, especially if combined with the expectation that the value of truthfulness must speak to us in the form of a simple, exceptionless law (something that, on Williams's view, pragmatic genealogical reflection itself gives us reason to doubt), it leads to the idea that everyone—including the villain at the door—always deserves the truth (2002, 122).

The needs visible in a model not yet informed by history also only call for truthfulness about the *immediate temporal environment*, but not about the *distant past*. And indeed, before Thucydides, there was no 'objective conception' (2002, 55) of the distant past. It is only at a particular historical juncture that the restricted conception of truthfulness about the past was replaced by one that stretches

---

[24] Williams says comparatively little about the rise of scientific truth, although at one point he had planned to include a section on the topic. He then abandoned that plan on the grounds that it was 'such a big subject and one so thoroughly treated by others' (Baker 2002).

[25] This is the striving of science for what Williams called *the absolute conception of reality*: 'a conception of the world as it is independently of the peculiarities of any observers' (2005b, 226). See also Williams (2006d, 185–9; 2011, ch. 8; 2014e, 323–4) as well as Moore (1997) for an elaboration of the idea.

further back in time. Nevertheless, this development is dimly anticipated in the state-of-nature model, because the ideas offered in that model—of something's being 'earlier than' something else, or of 'a day that people can remember'—suffice to put the objective conception of the past within reach (via ideas such as 'days earlier than any one of us can remember'; although there are of course difficulties in determining just when something is, in principle, 'within reach').

(c) *Autonomous Elaborations*: Third, the prototype can be elaborated by history in ways that are in no way anticipated by the state-of-nature model and entirely unpredictable from the armchair. These elaborations reflect the *autonomy* of the historical process relative to the abstract framework laid out in the state of nature (2002, 40, 173). As François Truffault quipped, life has more imagination than we do. A look at history reveals, for instance, that the virtue of Sincerity provided the material for the rise of notions of personal authenticity in the eighteenth century. Very roughly, it was no longer enough to say what one took to be true at a given time; the expectation now was that one should be authentic, that what one said should be not just what one believed, but an expression of who one *really* was. Such an elaboration does not answer to the demands initially set out in the state-of-nature model; nor does it realize a development that is already prepared in that model. Rather, it is a 'secondary elaboration' (1997a, 32); a product of culture that 'cannot be seen as a development of human needs, concerns, and interests which was inevitable, or even particularly probable' (2002, 172). Scientific inquiry in its purest forms—in which it is most independent from practical and economic interests—can also be viewed as having been elaborated beyond the specifications necessary from the point of view of the demands initially set out in the state of nature (1997a, 31–2). But this is not to say that these elaborations do not serve more local needs. Williams acknowledges this when he writes that there is no *one* reason to value truthfulness (2002, 263–5), and his genealogy's involvement in history serves to do justice to that fact. The state-of-nature considerations indicate widely shared instrumental reasons to value truthfulness intrinsically, but the later parts of the book point to other reasons that are more closely tied up with our history and our particular cultural situation. The achievement of other goods is instrumentally dependent on the cultivation of some measure of truthfulness: a sense of freedom in scientific inquiry (2002, ch. 6, §4), for instance, but also the prevention of tyranny in politics (2002, 207–8), Millian self-development along with political liberty as expressed in freedom of speech (2002, ch. 9, §2), and distributive justice (2002, ch. 9, §4).[26] But he also highlights a range of local needs that truthfulness serves. In particular, if we consider truthfulness from a socio-historical position marked by liberalism, we can be seen to have a local need for truthfulness *about political history*. Liberalism,

---

[26] See also Williams (2005k).

of the kind concerned to secure for the less powerful some measure of freedom from abuse of power and control, *specially needs* truthful history (2002, 263–5). This may seem counterintuitive, given that truthful history tends to soak in suspicion the flattering narratives of progress that liberalism tells about itself. But for Williams, liberalism is most minimally and most powerfully rationalized by fear. In contrast to the Lockean liberalism of natural rights or the Millian liberalism of personal development, the *liberalism of fear* (as Williams calls it following Judith Shklar) is animated by historical awareness of past atrocities, and in particular by the sense, 'amply justified on every page of political history', that 'some agents of government will behave lawlessly and brutally in small or big ways most of the time unless they are prevented from doing so' (Shklar 1989, 28).[27] Liberals accordingly have a special reason to value truthful political history over myth.

To sum up: the prototype presented in the state of nature is subject to three types of elaboration: (a) specification, which is both anticipated and necessitated by the model in its initial configuration; (b) extension, which is anticipated but not necessitated by the model; and (c) autonomous elaboration, which is neither anticipated nor necessitated by the model.

7.3 Reading Williams as a Cambridge Pragmatist

This reconstruction of Williams's genealogy suggests that although it is nominally directed against those he sometimes calls 'the pragmatists' (2002, 59–60, 128–30), it is illuminatingly described as a pragmatic genealogy and as a compelling example of Cambridge pragmatism in particular. This means that Williams has more in common with his declared opponents than he admits, but it does not mean that we lose the distinction between the position he advocates and the one he attacks. If we distinguish *truth-denying* from *truth-affirming* pragmatists,[28] we can view Williams as arguing against the former (as represented by Rorty), but on grounds that place him squarely among the latter—not merely because he affirms truth, but because his approach to truth is pragmatist in all but name. A truth-affirming pragmatist strives 'to illuminate the concept of truth by considering its linkages with inquiry, assertion, and the acquisition of belief' (Misak 2016, 28).[29] Much effort has recently gone into highlighting the more truth-affirming strands of pragmatism. As Cheryl Misak shows in her *Cambridge Pragmatism* (2016), these lead from C. S. Peirce in Cambridge, Massachusetts, to C. K. Ogden in Cambridge, England, who knew about Peirce thanks to the independent scholar

---

[27] Williams (2005f) elaborates on his conception of the liberalism of fear.
[28] Misak (2016, ix) and M. Williams (2013, 129; 2016, 223).
[29] See Wiggins (2002b, 316) and Misak (2015, 264).

Lady Victoria Welby. Ogden introduced the young Frank Ramsey to Peirce's work, and Ramsey of course went on to influence Wittgenstein and a long line of Cambridge philosophers—including E. J. Craig, whose methodology is openly pragmatist, and Williams himself, who preceded Craig as Knightbridge professor at Cambridge and falls squarely into the tradition of Cambridge pragmatism.[30] Indeed, his exploration of how pragmatist insights can be enriched with a dusting of history makes him a particularly interesting exponent of Cambridge pragmatism.

In line with the idea that the best understanding of philosophy's subject matter is agent-centred, Williams elucidates truth in terms of the various 'states and activities associated with the truth' (2002, 7). He eschews questions such as what truth *is*, asking instead after the role of truth in various human dealings, such as language learning, believing, asserting, inquiring, communicating, and cooperating. Part of what licenses the application of the label to Williams, then, is that he elucidates truth in terms of what *we do* with truth: the human dealings expressive of truthfulness. What also licenses the label's application, however, is that Williams raises a pragmatist question in the reverse direction, asking what truthfulness *does for us*. Most basically, truthfulness 'gets its point ultimately from the human interest, individual and collective, in gaining and sharing true information' (2002, 126). But as he then goes on to show, truthfulness does a lot more for us, and what *we*—as opposed to the agents in the state of nature—would lose if we lacked it goes beyond a loss of efficiency in gaining and sharing information. For all these reasons, his genealogy makes for a paradigm example of Cambridge pragmatism at its best.

It is true that Williams harbours a deep suspicion of labels and -isms, so that belabouring the question of how to label his genealogy may seem to run counter to his own conviction that labels often stand in the way of truthful inquiry—a point Williams likes to illustrate with an anecdote of Ryle's, who, after giving a lecture in Germany, saw a student come up to him say: 'I was very impressed by your lecture and would like to join your school. Unfortunately, I am a Kantian' (Williams 1995d, 186; 2007, 130). In fact, Williams's own use of the label 'pragmatism' to pick out his target itself illustrates the disvalue of labels, for it misleadingly obscures the pragmatism pervading his own enterprise. It is in correcting for this that the value of labelling Williams a Cambridge pragmatist lies—it derives its value not from the importance attaching to labels, but from the disvalue of attaching importance to labels. Once we recognize the Cambridge pragmatism in *Truth and Truthfulness*, we are in a position to discern a similar pragmatism in other works of Williams's: in his discussion of the concept of knowledge (2005b, ch. 2), of obligation (2006h, 73; 2011, 202–8), of legitimacy (2005c, e, i, j), of

---

[30] Misak (2016) traces these under-appreciated strands of pragmatism from Peirce to Wittgenstein in particular. See also Misak and Price (2017).

responsibility (1993, ch. 3), of the moral/non-moral distinction (2001b, 66–70), of the political value of liberty (2005a, c), and of the demand for ethical theory (1981a; 2005f, 54; 2011), for instance.[31]

Recognizing the Cambridge pragmatism in *Truth and Truthfulness* also enables us to see it as a direct response to the Rortyan question of why we should value the truth intrinsically. Williams's answer is that we should continue to value the truth intrinsically, and do so across a broad range of subject matters including politics and history, because we *need to* if we are not to lose much that is important to us, from the motivations fuelling liberalism and our achievements in securing political liberty and understanding history down to the very fabric of epistemic cooperation in society. Williams's genealogy both explains and vindicates our intrinsic valuing of the truth by revealing functionality where we did not necessarily expect it. It offers a perspicuous representation of that value's instrumental relation even to very basic human needs: if we only share the need for information, we share the need to value the truth intrinsically and to be able to do so.

Finally, it is worth noting that Williams's brand of pragmatism differs from the more 'thoroughgoing pragmatism' about the value of truth advocated, for example, by Stephen Stich (1990, 21). On Stich's view, 'all cognitive value', including that of truth, 'is instrumental or pragmatic', and 'there are no intrinsic, uniquely cognitive values' (1990, 21). The core of Stich's argument is that for a belief to be true is for it to be mapped to a true proposition by one among countless possible mapping functions, namely the function that happens to be sanctioned by our intuitions. But since there is nothing special or important about that intuitive mapping function—it is simply the 'idiosyncratic hodgepodge' (1990, 120) that our biological and cultural history disposes us to operate by— there is nothing special or important about truth.[32] This makes the intrinsic value of truth seem like 'a curious, culturally local value, on a par with finding intrinsic value in the cultural practices of one's ethnic group' (1990, 24). Williams's genealogy dispels any impression that the intrinsic value of truth is curious or parochial. Just as Stich tries to weaken our confidence in the intrinsic value of truth by presenting it as an arbitrary and idiosyncratic fetish, Williams tries to strengthen our confidence by showing that there is nothing mysterious or surprising about it: any social creatures in need of information would be bound to come to value the truth intrinsically, and it makes good practical sense for them to

---

[31] See Cueni and Queloz (2021).

[32] See Stich (1990, 23–4, and esp. ch. 5). The first step of Stich's argument, that our intuitive mapping function may in fact be suboptimal given our purposes, is one that Williams's genealogy is neutral towards. What Williams's genealogy blocks is the inference Stich then draws from 'the realization that the function is an idiosyncratic hodgepodge bequeathed to us by our cultural and/or biological heritage' to the conclusion that we should 'not value true beliefs intrinsically' (1990, 120). A more recent attempt to throw doubt on the value of true belief is Hazlett (2013). But as Hazlett himself acknowledges (2013, 144–5), his critique is orthogonal to Williams's argument, as it concerns not the value *of the practice of valuing* true beliefs, but rather the value of true beliefs.

do so. While Stich, like Rorty, renders it tempting to think that our purposes might be better served by taking a more instrumental view of the value of truth, Williams vividly shows that it would be a ruinous mistake to think that we could have the instrumental value without the intrinsic value. By refuting the pragmatist deniers of the intrinsic value of truth on pragmatic grounds, Williams proves them wrong by their own lights.

7.4 McGinn's Three Challenges and Self-Effacing Functionality

With this understanding of Williams's genealogy in place, we are now in a position to examine the reach of its vindicatory power, its relation to indirect utilitarianism, and its aptitude for dealing with self-effacing functionality. A helpful way into these questions is the critique of Williams's method advanced by Colin McGinn, which I shall use as a foil. McGinn not only highlights the resemblance of Williams's method to just the kind of utilitarianism of which Williams was a critic, but more generally calls into question this kind of use of genealogy. We learn much about the method by examining how it stands up to McGinn's critique.

According to McGinn, Williams's genealogy is vulnerable to three challenges:

(1) *The No-Intrinsic-Value Challenge*: 'showing the function that a virtue serves can only give it instrumental value, not intrinsic value...Since Williams insists, rightly, that truthfulness has an intrinsic value...his functional story fails, by his own standards, to capture that intrinsic value; so it does nothing, really, to vindicate the intrinsic value of truthfulness'.

(2) *The Utilitarianism Challenge*: 'the functional account looks like a thinly disguised form of utilitarianism, an argument to the effect that truthfulness is good because it increases the general level of human wellbeing'.

(3) *The Redundancy Challenge*: 'once this is seen the genealogy itself becomes theoretically redundant' (McGinn 2003, §1).[33]

McGinn's first challenge, that instrumental considerations in favour of truthfulness do nothing to vindicate its intrinsic value, renders acute the question of just what Williams's genealogy aims to achieve. The answer can be articulated by drawing once more on the triad of negative, naturalistic, and pragmatic vindication introduced in Chapter 4: Williams offers a *negative* vindication of truthfulness

---

[33] For a response to McGinn's challenges that differs from the one I go on to sketch, see Thomas (2008).

insofar as his genealogy does not excavate anything to suggest our endorsement of truthfulness to be radically self-deceived, thus clearing truthfulness of suspicion and marking it out as stable under reflection; and he offers a *naturalistic* vindication of truthfulness insofar as the genealogy enables us to make sense of truthfulness in terms of the *rest* of nature, in particular in terms of basic needs of cooperation and communication. Both of these vindications ensure that we can remain confident in truthfulness

> in the sense that we can understand it and at the same time respect it, support it and live within it. We can also urge it against alternative creeds whose own self-understandings (as divine revelations, for instance) are themselves not going to survive a genealogical inquiry.   (Williams 2014g, 410)

Yet Williams also offers a *pragmatic* vindication of truthfulness. His genealogy reverse-engineers the point of truthfulness, a point it is shown to possess relative to needs ranging from socio-historically local needs to needs so basic that they would be at work in anything recognizable as a human society. And as he urges against the deniers, that is very much a reason not to give it up.

It is here that McGinn's first challenge gets a grip. A pragmatic vindication may be fine as far as it goes, but it does not go far enough; showing what function something serves can give it instrumental value, but not intrinsic value; hence, the genealogy fails to vindicate the idea that truthfulness is intrinsically valuable. In fact, Williams himself acknowledges the force of this argument when he admits that the considerations he advances only show that we need to *treat* truthfulness as having intrinsic value—they do not and cannot show that truthfulness in fact has intrinsic value (2002, 90).

But what McGinn fails to see is, first, that Williams is not trying to *give* truthfulness intrinsic value—he is trying to vindicate the intrinsic value it *already has*; and second, that while it is true that the genealogy can only show that we need to *treat* truthfulness as an intrinsic good, it does not follow that this amounts to mere pretence. If pragmatic considerations show that regarding truthfulness as an intrinsic good is instrumentally necessary, the question is whether we *can* and *do* so regard it; if the answer is yes, we shall have been vindicated in regarding truthfulness as an intrinsic good. As we saw, Williams takes fulfilment of what we called *The Practical Necessity Condition* and *The Sense-Making Condition* to be sufficient, albeit not necessary, for something to be an intrinsic good. Merely seeing the benefits of valuing something intrinsically is not enough actually to internalize that value. But if both conditions hold,

> we have not simply adopted an illusion or a pretence of there being an intrinsic good. In fact, if these conditions hold, that would be a very odd thing to say, implying as it does that there is something further which would count as its *really*

being an intrinsic good, of which these conditions offer only a surrogate or mock-up. If the conditions are satisfied, then we shall have *constructed* an intrinsic good.   (Williams 2002, 92)

Williams need allow no sense to the term 'intrinsic good' except one in which it is true that Accuracy and Sincerity are intrinsic goods, because the two relevant conditions are satisfied. If these conditions are satisfied, we are not *pretending* that Accuracy and Sincerity are intrinsic goods.[34] That, as Williams points out, 'would be something else: it would be to pretend that we needed them, or to pretend that we could make sense of them as intrinsic goods, and neither of these things need be so' (2006f, 137).

In order better to grasp Williams's conception of intrinsic value, it helps to step back from his genealogy for a moment to consider the conception of intrinsic value in answer to which Williams presents his own, which is Christine Korsgaard's. Korsgaard maintains that 'intrinsic' goodness contrasts not with 'instrumental' but with 'extrinsic' goodness, while 'instrumental' is properly contrasted with 'final' goodness (1996, 249). 'Intrinsic goodness' as it occurs in the intrinsic/extrinsic contrast is a matter of how we *explain* goodness, and an intrinsic good is therefore one whose goodness is *self-explanatory*.

Distancing himself from this view, Williams urges us to 'give up the unrewarding idea of intrinsic goodness being self-explanatory' (2006f, 136). He proposes instead to recast intrinsic goodness in the same terms as the contrast between final and instrumental, namely as a contrast between *ways in which we value* things. Williams proposes to explain the notion of *being* an intrinsic good in terms of people *valuing it as* an intrinsic good, just as we explain being an end in terms of someone pursuing it as an end. This is not meant to erase the distinction between some agent's valuing something as an intrinsic good and its actually being an intrinsic good: we can still say, with Aristotle, that an intrinsic good is what a wise or rational person would value as an intrinsic good (Williams 2006f, 122). Where Williams is keen to break with the philosophical tradition from Plato through G. E. Moore to Korsgaard, however, is with regard to technical understandings of intrinsic goodness as something self-explanatory, enshrined in the intrinsic properties of an object and intelligible even when that object is considered *in vacuo*.

The fundamental contrast then becomes that between *valuing intrinsically* and *valuing derivatively*. Valuing as an end (as a final good) and valuing as a means (as an instrumental good) are applications of this general contrast. Some things are valued derivatively without being valued only as means—one may value going to the concert as a way of having a good evening, but this does not mean that one values it only as an instrumental good (2006f, 122–3). Similarly, some things, like

---

[34] See also Thomas (2008, 358–60; 2012, 155–7) for a defence of Williams on this point.

natural beauty and works of art, are valued intrinsically without being valued as ends, at least if valuing as an end is understood narrowly as entailing that someone does or would *pursue* it as an end (2006f, 135–6). Consequently, Williams's proposal does not reduce all intrinsic value to practical attitudes like wanting, pursuing, or trying to bring about; but it does entail that intrinsic value should be capable of being 'expressed in practical attitudes', for it 'would be an empty thing if it could not be expressed in such attitudes, even if there are some circumstances in which it can be expressed in no more than wishing' (2006f, 136).

The notion of intrinsic value that informs *Truth and Truthfulness* is thus distinctly non-Platonic: it is free of any demand that intrinsic goodness be self-explanatory, enshrined in intrinsic properties, or intelligible in isolation from everything else. And yet the book does owe something to Plato. This is a debt that Williams acknowledges in his essays on the *Republic*'s discussion of the intrinsic goodness of justice. As Williams writes there, 'perhaps the lesson of Glaucon's argument is just this, that precisely because we need justice as an instrument we need to admire it for its own sake; and that what we need to do is to learn how to do this, while not forgetting why we are doing so' (2006e, 107). In his explanation of intrinsic value, Williams takes himself to generalize Plato's schema for the explanation of justice:

> when we considered Plato's account of the final good of justice, we found that (relative to his own assumptions) he was able to explain it in terms of our needing to treat justice as a final good, and our being able to make sense of our doing so. My last suggestion is that we should extend this schema to intrinsic goodness in general. We give up the unrewarding idea of intrinsic goodness being self-explanatory. We say that something is intrinsically good if we need to value it as intrinsically good, and we can make sense of our doing so. . . . If we can make sense of trustworthiness in such terms, then we shall have *constructed* an intrinsic good.   (Williams 2006f, 136–7)

We recognize these two conditions as The Practical Necessity Condition and The Sense-Making Condition, and the line about constructing an intrinsic good is familiar from *Truth and Truthfulness*. But who exactly is supposed to be doing the constructing? Humanity in the course of its history? The genealogist in telling the story? The audience in understanding it? A later passage is clearer about who 'we' is:

> We have seen how much Plato achieves, in his own terms, by asking the questions 'do we need to value justice as a final good?' and 'can we make sense of our doing so?' If we use 'intrinsic' in the new and broader way, Plato will be asking these questions about justice as intrinsically good, and in answering them, he will have constructed an intrinsic good.   (Williams 2006f, 137)

This suggests that the one doing the constructing is the genealogist. But out of what is something being constructed?

At this point, Williams would be the first to insist that we must remember to ask, of any genealogy, *to whom* it is addressed. Who needs to hear it? Whom does it aim to convince? In the case of Williams's own genealogy, the target audience are those who harbour doubts about the intrinsic value of truth, either in a certain area or across the board. The target audience is not some alien intelligence; nor is it that bugbear of philosophy, the amoralist who shares no moral values with the genealogist and who needs to be reasoned into a way of life. The real people to whom Williams's genealogy is addressed already participate in a way of life in which the value of truth has a long and rich history, but their confidence in that value has been undermined by suspicions—that the value of truth cannot really be made sense of in naturalistic terms, perhaps, or that there is no point in valuing the truth when everything comes down to power. It is for such an audience, out of the material that the genealogist shares with this audience, that an intrinsic good needs to be constructed.

Williams constructs an intrinsic good not by giving it a value it did not have before, but in something like the way Wittgenstein assembles reminders: in telling us his genealogical story, Williams shows us that we do in fact already possess both the material to make sense of truthfulness *as* an intrinsic good, because we can think reflectively about truthfulness, and relate it to other things we value, in a way that need not reduce it to a mere device for maximizing utility or solving game-theoretical problems, and the material to make sense of *our valuing* it as an intrinsic good, because it makes sense for people with our needs to do so—we have not simply stumbled or been misled or deceived into doing so. He thus constructs an intrinsic good in the sense that he puts together, in the right way, all the pieces required to present truthfulness to reflection as really being an intrinsic good that merits our confidence. For the genealogy to succeed, it is enough for the pieces to be capable of commanding the confidence of his audience. They need not be intelligible in advance of any human valuation.

The core of the answer to McGinn's first challenge, then, is that functionality is not supposed to *give* truthfulness intrinsic value, but to help us make sense of it and vindicate it as the intrinsic good it *already is*. Williams's vindication of truthfulness is not addressed to the amoralist, and it does not attempt to vindicate truthfulness from some Archimedean point outside the ethical life. It is addressed to people who possess the resources to make sense of it as a value, and who are considering what reasons they have for continuing to value truthfulness. If this is the task, Williams fulfils it. For he only has to show that given a commitment to certain needs, we have reason to value truthfulness as an intrinsic good, and can in fact make sense of it as an intrinsic good; and he achieves this, for he derives the need to value truthfulness as an intrinsic good from a series of generic and local needs, and traces out the connections of truthfulness to a variety of others things

that we value. In doing so, he constructs an intrinsic good—not in the sense of creating it *ex nihilo*, but in the sense of highlighting, drawing together, and revealing connections between considerations that are, for us, already there. The genealogy offers reasons for those who are capable of seeing truthfulness as intrinsically valuable to continue to do so, thereby helping to continually create a community held together by this value. The genealogy is not an instrument of conversion. But by promoting self-understanding, it props up the confidence of those who are, in some measure, already disposed and equipped to value truthfulness for its own sake.

In vindicating truthfulness, Williams also pursues a wider concern: to show that, in a seeming paradox, 'intrinsic values…have their uses' (2002, 127). In an atmosphere dominated by what Williams perceives as undue 'scientism' (2006d), many are suspicious of intrinsic values, and are attracted by theories like utilitarianism partly because these promise to make sense of so much of the world in instrumental terms, which can seem like the only naturalistically intelligible form of value. Any such theory has to be grounded in some intrinsic values; but at least they are kept to a minimum. Williams, by contrast, makes us comfortable with intrinsic values—partly by showing that they can be made sense of without deteriorating into pretence, and partly by showing that instrumental considerations themselves call for intrinsic valuing. This exemplifies what he elsewhere describes as an important contribution that philosophy can make, namely 'to liberate', in particular 'by suggesting to people that they really have a right to some conception, which has been condemned by a simple or restrictive notion of how we may reasonably think' (1995b, 233–4).

Let us now turn to McGinn's second challenge—that Williams's functional story constitutes a thinly disguised form of utilitarianism. This need not be problematic in itself, but it would be a problem by Williams's own lights, since the form of utilitarianism his vindication most closely resembles is the indirect utilitarianism he himself impugned as unstable under reflection.[35] Turning Williams's objection back on himself, we might say that the attempt to justify the disposition to value truthfulness intrinsically on purely instrumental grounds leads to a tension between the *spirit being justified* and the *spirit that justifies*. Under reflection, such a structure is bound to unravel. This is not for the reason we encountered above—that we need to be able to make sense of the spirit being justified from the inside. The reason it is unstable under reflection is that it tries to combine the following two incompatible thoughts:

(1)   Truthfulness is intrinsically valuable.

(2)   Truthfulness is only instrumentally valuable.

---

[35]   See Smart and Williams (1973) and Williams (1995c, 2002, 90–1; 2011, ch. 6).

Indirect utilitarianism typically tries to relieve the tension by appealing to a distinction between theory and practice: we might think (2) in what Joseph Butler called the 'cool hour' of reflection, yet in the thick of the action, we focus firmly on (1). But on Williams's own account, the distinction possesses 'no saving power' (1995c, 165). How is his own story any different?

One difference is that because Williams does not share the utilitarian's commitment to there being only one *really* intrinsically valuable type of thing ('well-being' in McGinn's version), Williams is free to maintain that the two thoughts being combined are really these:

(1)   Truthfulness is intrinsically valuable.
(2')   Truthfulness is instrumentally valuable.

On the indirect utilitarian account, the recognition of truthfulness' instrumental value is achieved at the cost of its intrinsic value. But with (2'), there is no longer a contradiction, since the instrumental value ascribed is not *exclusive*. The recognition of instrumental value coexists harmoniously with that of intrinsic value.

Another difference is that while Williams offers a two-level view that is *structurally* similar to indirect utilitarianism, the repartition of the *justificatory weight* across Williams's two levels, (1) and (2'), is very different from its repartition across the two levels of the indirect utilitarian, (1) and (2). The utilitarian can agree that people's motivations in being truthful should be that truthfulness is a good thing in its own right, but what *really justifies* thinking in this way is the consideration, offered at the more reflective second level, that this is ultimately more conducive to well-being. For Williams, by contrast, the bulk of the justificatory weight lies on the first level: the fact that truthfulness is intrinsically valuable carries *more* authority than the fact that things go better if we think this way. The latter fact yields in the first instance an *explanation* rather than a justification, even if a vindicatory one.

The propensity to conceive of intrinsic and instrumental value as mutually exclusive is not specific to utilitarianism. It also manifests itself in the view that there is nothing more to be said about a good once we have recognized its intrinsic value—it is valuable, and that is all. On this view, an explanation that exhibits truthfulness' instrumental relations to other, less refined values will appear to besmear it—to imply that truthfulness has no intrinsic value (really). Once again, however, this appears so only if we conceive of the distinction between intrinsic and instrumental as an exclusive one: either we make sense of a value on its own terms, which we treat as irreducible, or we view it purely as a means to an end and reduce its value to that of something else.[36] But this sets up a false dichotomy.

---

[36] See Williams (1997a, 24; 2002, 90).

Once one rids oneself of the exclusive conception and understands intrinsic value as Williams suggests, it becomes evident that myriad things unite both aspects, and sometimes even possess one aspect in virtue of the other: Shelly Kagan (1998, 283–5) shows that something can possess intrinsic value in virtue of its instrumental value; Williams shows that something can possess instrumental value in virtue of its intrinsic value.

Some have nevertheless doubted that we can be 'aware *at the same time* of the intrinsic quality of a value and of its instrumental quality' (Hartmann and Saar 2004, 392) in the way Williams suggests.[37] Yet upon closer inspection, his genealogy side-steps this question anyway—the instrumental value is not located at the same level as the intrinsic value: while what is intrinsically valuable is truthfulness itself, it is the *valuing* of truthfulness as intrinsically valuable that is instrumentally valuable.

In fact, while truthfulness needs to be sustained by non-instrumental motives, awareness of the instrumental or functional motives for engaging in truthfulness is not required. No instrumental valuing of truthfulness is needed to reap its instrumental value. (2') might, however, be read this way, so to pre-empt this reading, the genealogy is best represented as issuing in the following two beliefs:

(1) Truthfulness is intrinsically valuable.
(2") The attitude expressed in (1) is instrumentally related to certain humans needs if widely shared and known to be shared.

It is only *the fact that* we value truthfulness intrinsically that is shown to have instrumental value, even if we never value truthfulness instrumentally. The genealogy relates 'a value which gives us some reasons for action to other reasons for action which...we have "anyway"' (Williams 2000, 160)—but while the relation *between* these reasons for action is an instrumental relation, the reasons themselves are not instrumental reasons. Functionality is not part of the *content* of the motives, but is *possessed* by them.[38] Functionality helps explain why we value what we value, but it is what we value that motivates and justifies what we do. This puts even more distance between Williams's genealogy and the problem of reflective instability.

This brings us, finally, to McGinn's third challenge: once we recognize the allegedly 'utilitarian' nature of the genealogy, McGinn argues, a functional account

---

[37] Hartmann and Saar take Williams to be reacting to what they perceive as a problem by integrating 'an instrumental component...into the very definition of intrinsicness', i.e. maintaining that 'intrinsic' just means 'non-egocentric' or 'not merely self-interested' (Hartmann and Saar 2004, 393). I offer a different reading.

[38] Craig's reading of Williams, by contrast, seems to retain the idea of 'functional motivations' (2007, 200).

can be given directly, without the developmental narrative. But this is precisely what Williams is at pains to deny: 'In relation to institutions, practices, expectations, and values that actually exist, of justice, promise-keeping, truthfulness, and so on, functional accounts are simply false'; it is 'just not true', he continues, 'that the dispositions of truthfulness that we have, or that anyone else has had, can be adequately explained in functional terms' (2002, 34–5), because these dispositions essentially involve non-functional motives, i.e. motives that make no reference to the practice's functionality. An explanation of our actual practice of truthfulness in functional terms would be inadequate in three respects: it would distort our understanding of our own practice by representing it as being primarily animated by functional motives when it clearly is not; it would discourage us from seeing what is peculiar to our own particular socio-historical elaboration of the practice; and it would fail to adequately explain *why* the practice is not primarily animated by functional motives.

There is a place for explanation in purely functional terms, but where self-effacingly functional practices like truthfulness are concerned, that place is not in a description of actual practices, either those we now have or those in our distant evolutionary past. Its place is in the state-of-nature fiction:

> The fiction is uniquely useful because—so far from confusing genuine history and fiction—it enables us to keep count of what is history and what is abstraction, and it helps us to avoid two errors. One is that of going straight to our actual society with the apparatus of functional explanation; this would distort our understanding of our own cultural situation, debar us from seeing what is peculiar to it as opposed to others, and lead us to a stupid reductionism. The other error is to construct pictures of very early societies on the basis of functional ideas and suppose that this was actual hominid prehistory. Genealogy keeps historical fact and functionalist abstraction in their places.
>
> (Williams 2002, 35)

This passage clearly shows that Williams subscribes to a dynamic model interpretation rather than to an actualist interpretation of genealogy: he thinks that there is a place for the functionalist or instrumentalist approach to concepts which makes sense of them as growing out of needs, but because it initially encourages a crude instrumentalism that ties human thought more closely to needs than it actually is or can ever have been, its place is in the realm of idealization or abstraction rather than in descriptions of real societies or in conjectures about prehistory. To equate our actual conceptual practice with the purely functional prototype in the state-of-nature fiction by assuming that this is all there is to our practice would be to distort our understanding of our situation by debarring us from seeing, first, the value we attach to truthfulness for its own sake; second, the further respects in which our practice of truthfulness differs from the generic

prototype in virtue of its historical elaboration; and third, the functional reasons for which we value truthfulness for its own sake.

If there is something functional about a practice that is not understood in functional terms, an adequate explanation of this will be one that accounts not only for the functionality of the practice, but also for the fact that it is not understood in functional terms. Williams's genealogy provides such an explanation by showing that the functionality of the practice itself requires the practice to be understood in non-functional terms: the presence of non-instrumental motivations to be truthful is part of the functional requirements on the practice of truthfulness—it can be stable, and possess instrumental value, *only insofar* as it is driven by non-instrumental motives. To miss these non-functional motives is not only to miss what makes the practice functional, but also to miss the main justification—as opposed to explanation—for engaging in the practice, since the non-functional motives may—indeed, from a practical point of view, should—possess more authority than any functional motives there might be for having those non-functional motives. The self-effacing functionality of truthfulness is an important part of the reason why Williams resorts to the genealogical method (2002, 93): the method is uniquely suited to capturing the dynamics of self-effacing functionality, because it can show that intrinsic values have their uses without the insight into the use leading us to lose our grip on the intrinsic value.

Williams's genealogy can thus be defended against McGinn's challenges on all fronts. It does not give truthfulness intrinsic value, but rather vindicates truthfulness as an intrinsic value, and in the process helps us make sense of intrinsic values in naturalistic terms; it is not a form of indirect utilitarianism, but differs from it in just the way required to avoid the problem of instability under reflection; and its functionalism does not render its genealogical dimension redundant—on the contrary, the genealogical dimension helps explain in functional terms why truthfulness must outgrow mere functionality, so that functionality ends up being a property of our motives for being truthful without showing up as part of their content.

7.5 A Pessimism of Strength: Williams's Debt to Nietzsche

Before we leave Williams, it is worth considering the nature of his professed debt to Nietzsche (2002, 13), since the effort made in Chapter 5 to tease out the English genealogist in Nietzsche has revealed continuities that run much deeper than was apparent at the beginning of our investigation, when their respective genealogical methods seemed to share little more than a name.[39] In fact, their topics, aims,

---

[39] Clark (2015b) offers a complementary account of Williams's debt to Nietzsche that focuses on the respects in which the critique of the morality system in *Ethics and the Limits of Philosophy*

methods, and insights are remarkably alike. Nietzsche memorably characterized the broader aim of his genealogy of moral values as being to determine '*the value of these values*' (*GM*, P, §6), and we saw that already in his genealogy of the will to truth, he asks after 'the value of this will' (*TL*, §1). Williams no doubt self-consciously echoes Nietzsche when he characterizes 'the principal aim of all moral philosophy' as being that of achieving an 'ethical understanding of the ethical'—of 'truthfully understanding what our ethical values are and how they are related to our psychology, and making, in the light of that understanding, a valuation of those values' (1995a, 578).

Yet even at the more substantive level of how they pursue this shared aim, Nietzsche and Williams tell strikingly similar naturalistic genealogies of truthfulness. For one thing, both genealogies are imaginary, vindicatory, and revelatory of necessity before they are historical, critical, and revelatory of contingency. But the parallels run deeper. Both Nietzsche (*eKGWB*, 1885, 40[43]) and Williams (2002, 125) take it to follow from their genealogies that truthfulness must include resistance against deception from 'within', or what Williams discusses under the heading of 'self-deception and wishful thinking'. They also both conclude that truthfulness presupposes a belief in what one calls the 'stability of the person' (*eKGWB*, 1883, 24[19]) and the other calls the process of 'steadying the mind' (2002, 191). Both Nietzsche and Williams are also keen to rely on what they have identified as the proper remit of truthfulness to criticize forms of truthfulness that overstep that proper remit: they both deem a 'blind rage for collecting, a restless raking together of everything that has ever existed' (*UM*, II, §3) and 'terminally mindless fact-acquisition' (2002, 256) to be regrettable outgrowths of truth-seeking or Accuracy; Nietzsche warns against pushing science forward as quickly as possible on the grounds that this would destroy it, 'just as a hen perishes if it is compelled to lay eggs too quickly' (*UM*, II, §7), while Williams makes similar points about pushing inquisitiveness and suspicion too far (2002, 2–3, 15, 212, 301n44). On the side of truth-telling or Sincerity, they both deny (*contra* Kant) that a commitment to it entails that everyone equally deserves being told the truth (*eKGWB*, 1885, 40[43]; Williams 2002, 122), or (*contra* Rousseau) that it always implies complete disclosure (*eKGWB*, 1886, 7[6]; Williams 2002, 85, 109, ch. 8).

Of course, it has also emerged that Williams is more deeply indebted to Hume than he lets on. But although Williams harboured great admiration for Hume early in his career, he later came to take issue with what he called Hume's 'somewhat terminal degree of optimism' (1999, 256). This gnomic remark has

---

echoes Nietzsche's critique of Christian morality. For a contrasting discussion that emphasizes differences between Nietzsche and Williams, see Leiter (manuscript). For a reconstruction of Williams's critique of Nietzsche's ambition to formulate a criterion by which to select better values, see Queloz (forthcoming-a).

been developed in various directions,[40] but one thing it plausibly points to is Hume's sanguine confidence that our ethical concepts will prove stable under genealogical reflection. As we saw, Hume thought that our sense of morals 'must certainly acquire new force when, reflecting on itself, it approves of those principles, from whence it is derived, and finds nothing but what is great and good in its rise and origins' (*T*, 3.3.6.3).

This sunny picture sharply contrasts with the darker one painted by Nietzsche, who is more pessimistic about our prospects of being left with a sufficient number of ethical concepts to sustain a meaningful life once we subject them to truthful scrutiny. In this respect, Williams is closer to Nietzsche than to Hume. Williams agrees with Nietzsche that there is a 'need to be sceptical' (2014c) towards our ethical ideas—though not in the indiscriminately subversive fashion of positivists and error theorists whose scepticism derives from timeless and highly general considerations; Williams believes that there is both more to be feared and more to be learned from a scepticism that is partial and historically situated—a scepticism that casts suspicion on particular tracts of our ethical thought, either because of their psychological origins, or because in our historical situation they 'may no longer do what they once did for us' (2014c, 317). While this kind of scepticism 'has done only too well in some historical and literary studies', Williams observes, it still needs 'to take a black look at the received pieties of much moral philosophy' (2014c, 318).

Williams thus follows Nietzsche in thinking that there is a real challenge involved in achieving a form of confidence in our concepts that is not grounded in comforting myths, illusions, or blind dogmatism, but in a truthful recognition of how these concepts came to be ours. As he concludes in *Ethics and the Limits of Philosophy*:

> One question we have to answer is how people, or enough people, can come to possess a practical confidence that, particularly granted both the need for reflection and its pervasive presence in our world, will come from strength and not from the weakness of self-deception and dogmatism. (Confidence is not the same as optimism; it could rest on what Nietzsche called the pessimism of strength.)
>
> (2011, 190)

Pessimism in Nietzsche's sense involves open-eyed awareness of the pervasiveness of suffering, of its entanglement with humanity's greatest achievements, of human exposure to chance, uncertainty, and sudden blows of fate; but perhaps most importantly, it involves awareness of the rational contingency of it all: much happens without a deeper reason or a higher purpose, and there is no credible

---

[40] See Greco (2007), Sagar (2014), Russell (2019), and Blackburn (2019).

redemptive narrative that will recast this contingency as a form of necessity, or guarantee that it will all be worth it.[41] When such pessimism leads to a negation of life—be it in the form of a death-wish or of a denial of reality through escape into myth and illusion—it is expressive of weakness in Nietzsche's eyes.[42] He encourages us to aim instead for a 'pessimism of *strength*' (*eKGWB*, 1887, 10[21]): a state in which one truthfully recognizes what reality is like in all these respects, but finds oneself capable of accepting and affirming this reality instead of succumbing to the temptation to rationalize or deny it.[43] It is this call for a pessimism of strength that Williams echoes in the final words of *Truth of Truthfulness*, when he expresses the hope that 'the ways in which future people will come to make sense of things will enable them to see the truth and not be broken by it' (2002, 269).[44]

If Williams ends up echoing Nietzsche's call for a pessimism of strength, it is not least because he and Nietzsche respond to the same basic predicament, a predicament growing out of two convictions they both share: that we should aim for truthful self-understanding, but that our ideas are unlikely always to survive truthful scrutiny. Deeply committed to the idea that ethical thought, in particular, 'should stand up to reflection' and 'its institutions and practices should be capable of becoming transparent' (2011, 222), Williams is also inclined to agree with Nietzsche's diagnosis of the condition of modernity as 'one in which we, at once, have a morality which is seriously unstable under genealogical explanation; are committed (by that very morality, among other things) to transparency; and find very little to hand in the way of an alternative' (2000, 160). Like Nietzsche, Williams thinks that to 'recognize how we are placed in this respect is, if anything, an affirmation of strength', and to 'suppose that the values of truthfulness, reasonableness, and other such things that we prize or suppose ourselves to prize, are simply revealed to us or given to us by our nature, is not only a philosophical superstition, but a kind of weakness' (1995e, 148). Indeed, it is a weakness that ultimately threatens to leave us with no values at all. A Platonic sensibility might experience Williams's genealogy as subversive, because it traces the value of truth only to its instrumental origins in the contingent muck of human concerns. But in a truthful, reflective, scientifically advanced, and historically self-conscious society, that kind of sensibility must eventually find it hard to

---

[41] See Nietzsche (*GM*, II, §7, III, §28; *BT*, Self-Criticism, §1; *eKGWB*, 1885, 2[100]). See also Owen (2007, 18) and Reginster (2006).

[42] See Nietzsche (*eKGWB*, 1885, 2[100]).

[43] Williams discusses Nietzsche's pessimism in a number of other places (2006b, j, l; 2012). See also Jenkins (2006, ch. 8).

[44] Williams once remarked that the bleak final stanza of Matthew Arnold's 'Dover Beach' summed up his attitude to life (McMahan 2013, 19). The Nietzschean theme of finding the strength to affirm life in the face of 'very compelling true accounts of the world that could lead anyone to despair who did not hate humanity' (2002, 268) is echoed already in the quotation from Wallace Stevens—another avid reader of Nietzsche—that Williams chose as epigraph for *Ethics and the Limits of Philosophy*. Krishnan and Queloz (manuscript) situates Williams and this epigraph in the historical context of World War II and its aftermath. See also Krishnan (forthcoming) for a narrative history of philosophy in Oxford between 1900 and 1960.

maintain its confidence in *any* values. For if a metaphysical conception of values is intolerant of any genealogical connection with mundane practical concerns, truthful reflection on our values' origins is bound to prove indiscriminately subversive. And if the conception of values that makes Williams's genealogy of the value of truth look subversive is one that makes any truthful genealogy look subversive, it is the conception that should be called into question, and not the value of truth.[45] We need to make sense of the world in terms that help us to live in it, and that must include valuing things in it.

But while we can be confident that people 'will continue to make sense of the world in terms that help them to survive in it', the real question is 'how truthful those terms can be, and how far they can sustain the more ambitious ideals of truthfulness that we possess, together with institutions that both help to make those ideals effective and can themselves be sustained in knowledge of the truth' (2002, 268). This is the concern that drives Nietzsche and Williams to try and 'see how far the values of truth could be revalued, how they might be understood in a perspective quite different from the Platonic and Christian metaphysics which had provided their principal source in the West up to now' (2002, 18). Williams's genealogy shows us a way of truthfully affirming the intrinsic value of truth without the metaphysical conceptions of values of Platonic or Christian metaphysics—not just the minimal form of the value of truth that we are bound to cultivate as long as we communicate, but also 'the more courageous, intransigent, and socially effective forms' (2002, 269) into which it was elaborated in the course of its history.

It is no coincidence that the idea Williams chose for what he knew would be his last book should have been truthfulness. Truthfulness, he declared, is 'a first-order value of mine of the Nietzschean kind' (1999, 258). But there are also less personal reasons motivating the choice. Truthfulness is a thick concept—a concept that is both world-guided and action-guiding—and Williams was impressed by the fact that thick concepts offer us both more and less stability than thin ones like *good* and *right*. They offer us *more* stability in that they help stabilize the practice they are involved in by rendering judgements straightforwardly true (2011, 222); but they also offer us *less* stability in that they are particularly vulnerable to being unseated by reflection (1995d, 207); and for Williams, our 'major problem now' is that we risk being left with too few ethical ideas, and 'need to cherish as many as we can' (2011, 130). In showing that truthfulness proves stable under reflection, Williams not only defended an ethical idea, but demonstrated that there *are* thick concepts capable of surviving reflective scrutiny. And truthfulness is a particularly significant one, since it is a driving force of genealogical reflection itself. By demonstrating that truthfulness is capable of withstanding its own scrutiny,

---

[45] The argument that anyone has reason not to conceive of values in terms that render truthful genealogical inquiry invariably subversive is elaborated in Queloz and Cueni (2019).

Williams showed that genealogical reflection need not peck into the dust the tree that supports it. This puts genealogical reflection on a firm basis and invites further genealogical inquiry into which thick concepts we have reason to live by. It did not take long for the invitation to be taken up. Five years after the publication of *Truth and Truthfulness*, Miranda Fricker presented a genealogy of a third virtue of truth: the virtue of testimonial justice.

8

A Political and Ameliorative State of Nature

Miranda Fricker

Miranda Fricker's *Epistemic Injustice: Power and the Ethics of Knowing* (2007) has enjoyed far more attention than the genealogies of Craig and Williams it builds on. Yet surprisingly little of that attention has been directed towards Fricker's own genealogical narrative in that book, which has left many wondering what the genealogy is supposed to add to Fricker's case. Her genealogy is clearly a pragmatic genealogy, which uncovers the practical origins of what she calls 'the virtue of testimonial justice'. We exhibit this virtue when, in assessing the credibility of those who impart information to us, we reliably correct for the influence of prejudice. The virtue of testimonial justice fills a need to fend off 'testimonial injustice', which occurs when someone's word is not given the credibility it deserves. An example of testimonial injustice is the way in which, in Anthony Minghella's screenplay for *The Talented Mr Ripley*, the wealthy businessman Herbert Greenleaf reacts to the testimony of Marge Sherwood, the fiancée of his murdered son. When Marge voices her well-founded conviction that the murderer is the eponymous Mr Ripley, Herbert Greenleaf brushes her off with the line: 'Marge, there's female intuition, and then there are facts' (Minghella 2000, 130).

The insight that Fricker's genealogy pursues is that prejudices such as this one exert a counter-veridical pressure which interferes with the practice of information pooling envisaged by Craig and Williams. By abstracting entirely from social categorization in their state-of-nature model, Craig and Williams ignore the ways in which, in any human community, identification with certain groups and political relations between groups are bound to exert a distorting influence on the communal practice of acquiring and sharing truths. It is as if they considered the skills involved in archery without factoring in the wind. Once we factor in the confounding influence of social categorization, Fricker argues, we come to see that a third fundamental virtue of truth—the virtue of testimonial justice—is required to maintain one's aim at the truth in the face of gusty prejudice.

Since Fricker's genealogy is a seamless continuation of the project begun by Craig and Williams, much of the material in Chapters 6 and 7 also applies to her genealogy, and rather than repeat it, I shall focus on what is specific to Fricker's genealogy. Two features in particular stand out. First, her signature move is to de-idealize her state-of-nature model just enough to introduce *social and political categorization* into it. This brings into view the practical challenges that social and

political relations between groups raise for the practice of information pooling. Second, her use of the state-of-nature model is in important respects *ameliorative* rather than explanatory. All the genealogists we considered so far reverse-engineer the points of concepts or virtues that already figure prominently among the ideas we live by. Fricker's genealogy of testimonial justice is different: it is not about a deeply familiar and ubiquitous trait whose very ubiquity calls for explanation; it is about a trait that is as yet insufficiently familiar and not ubiquitous enough. Her genealogy issues in a model of epistemic division of labour that is offered as a corrective, indicating respects in which our conceptual practices fall short of serving some of our needs as well as they could. Fricker thus goes beyond her predecessors by introducing a political dimension into her state-of-nature model and by using that model to show how we can improve our practices. In her hands, more so even than in the hands of her predecessors, genealogy is conceptual reverse-engineering in the service of ameliorative conceptual engineering.

8.1 Good Recipients of Information

Let me begin by giving a slightly idiosyncratic reconstruction of Fricker's genealogy and pointing out where it branches off from its predecessors. We saw in Chapter 7 how Williams moves in his genealogy from the need for an epistemic division of labour to the need for good contributors to the pool of information—people who are able and willing to acquire new information and to pass it on to others. It is at this point that Fricker parts company with Williams. The need for an epistemic division of labour, Fricker suggests, entails not only the need for good contributors to the pool of information, but also the need for *good recipients* or, as Fricker calls them, 'hearers' of information—people who are able and willing to receive information through testimony. It seems to me that the notion of a *recipient* of information should not simply be identified with the Craigean notion of an *inquirer*, since that would make inquiry in one respect much easier and in another respect much harder than it can usefully be represented as being. It would be much easier—ludicrously easy—if being an inquirer as to whether p were sufficient to make one a recipient of the truth as to whether p. But it would also be much harder if being the recipient of a piece of information required one to stand to it in the relation of an inquirer, for it is one mark of an effective system of epistemic division of labour that people can gather and preserve information as to whether p without necessarily being themselves in the position of inquirers as to whether p. The communal availability of information is greatly increased by the fact that each of us picks up more of it than we intentionally seek out.

Nevertheless, the practical demands on good recipients of information overlap to a large extent with those on inquirers, and as Craig shows, an inquirer minimally needs a concept by which to pick out good informants on a given

matter—a concept tracking indicator properties that correlate with being right as to whether *p*. Now Fricker's guiding insight is that *being suitably sensitive* to the presence of these indicator properties is more work than it appears on Craig's telling. There is more to it than just possessing the concept of proto-knowledge, for wielding the concept properly requires resistance to a number of confounding factors.[1]

These confounding factors only come into view once we de-idealize the state-of-nature model a little further and factor *social categorization* into it, thereby introducing 'the relations of insider and outsider that human groups generate, and the relations of allegiance and enmity that naturally spring from them' (Fricker 2007, 115). Carl Schmitt rather sinisterly said that the fundamental political relation was that between friend and enemy, and in this sense of 'political', we might say that Fricker gives her state-of-nature model a political dimension.[2] As soon as we take social categorization into account, our state-of-nature community breaks up into groups with distinct identities.

But with social categorization, we also introduce the potential for *prejudice* into our state of nature. Fricker explores in rich detail how stereotypical images of social types prejudice us and influence our epistemically charged perception of others, i.e. our perception of them *as credible* to this or that degree. Through these mechanisms, the mere fact that someone belongs to a certain group—'ignorant outsiders, rivals out to trick one' (Fricker 2007, 116)—can dispose us to give them either *more*, or, in the case Fricker focuses on, *less* credibility than we would otherwise have given them. This pressure towards 'prejudicial stereotyping' (2007, 116) poses many problems, but the one that the genealogy sheds light on is the problem it poses for the epistemic division of labour. This division cannot function well if participants are insufficiently receptive to testimony, refusing to take on board information that is offered to them. This hampers the flow of information through the epistemic community.

But equally, though Fricker does not emphasize this much, recipients of information can be overly receptive by giving informants more credibility than they deserve. Fricker acknowledges that credibility attributions can be biased in the direction of excess as well as deficit, and that in localized contexts, this can be a problem (2007, 17–21): she gives the example of a doctor who is burdened by the fact that his patients expect him to advise them on medical matters which call for more specialist training than he possesses, and of a professor who is given so much

---

[1] This is the starting point of Elgin's (2019) critique of Craigean epistemology which we encountered in Chapter 6.

[2] Schmitt (1996, 26, 35). Williams (2005c, 77–8) incorporates this into a broader characterization of the political, which notably includes the idea that the political is focused on disagreement about what to do under political authority (paradigmatically, state power). But the notion of political authority, so central to Williams's political thought and to the Hobbesian tradition it stands in, does not yet figure in Fricker's state-of-nature model. In this sense, her state of nature is still pre-political.

benefit of the doubt by admiring colleagues that they effectively let her down by being insufficiently critical. On the whole, however, excess credibility will tend to work to the speaker's advantage, and Fricker concludes from this that we need not worry too much about it. But while the effects of excess credibility may well be net positive if we look only at the speaker, as Fricker's two examples do, there are other ways in which excess credibility might prove problematic.[3] Most basically, the division of epistemic labour cannot function well if recipients are overly receptive, gullibly accepting whatever is thrown their way, because this vitiates the pool of information by introducing falsehoods into it. Excess credibility also turns out to be problematic in further respects once we consider the connection— which it is very much in the spirit of Fricker's account to consider—between knowledge and power. It is a guiding idea of Fricker's book that a power deficit entrains a knowledge deficit in the sense that it leads to knowledge attributions being unjustly withheld, which in turn exacerbates the power deficit by giving one less of a say in the run of things. The flipside of this idea, however, is that to treat people as more knowledgeable than they really are is to grant them more power than they deserve. And this notably involves a loss in *freedom* on the part of the recipients who defer to the experts, a fact that becomes visible in the state-of-nature model as soon as we factor in even a primitive form of the need for freedom. Insofar as deference to experts can be reconciled with freedom at all, it will be on the back of the idea that the deference is warranted by the fact that the experts *merit* special credibility. By definition, *excess* credibility cannot be redeemed this way. Credibility excesses are thus likely to have not only epistemic costs, but also costs in freedom.

Whether we consider distortions of credibility judgements in both directions or only in the direction of deficit, however, it is clear that recipients of information need to resist prejudicial distortion—they need something that equips them 'to neutralize the impact of all sorts of prejudices endemic in the climate of testimonial exchange' (Fricker 2007, 96). A salient way of filling this need, Fricker argues, is to cultivate, alongside the proto-virtues of Accuracy and Sincerity, the proto-virtue of *Testimonial Justice*—the capitals again marking the fact that we are dealing with 'the abstracted virtue proper to the State of Nature scenario' (2007, 118). It is a virtue that can take two forms. It takes a *reflexive* form when its manifestation involves becoming aware of the likely impact of social identities— both one's own and that of the testifying party—and attempting to reverse and neutralize, through reasoning along the lines of 'Just because he's not one of us doesn't mean he's a liar/a fool' (2007, 116), any distortion in one's credibility judgements that may have been the result of prejudice. But the virtue can also take a spontaneous form when recipients of information perceive others through a

---

[3] For an influential discussion of the problem of credibility excess, see Medina (2011); see also Medina (2013, ch. 2) and Davis (2016).

'well-trained testimonial sensibility' (2007, 71) *from the start*, resisting rather than correcting the counter-veridical influence of social categorization. Either way, the point of the virtue of Testimonial Justice is to facilitate accurate judgements concerning the credibility of informants despite the distorting influence of prejudice (2007, 92). The virtue derives its point ultimately from the very same need for an epistemic division of labour from which Craig and Williams derive the need for the concept of proto-knowledge and the virtues of Accuracy and Sincerity. It is just that for the need for this third virtue of truth to come into view, we need to factor in, if only in its most schematic form, the fact that any epistemic division of labour will be complicated by the politics of 'us' and 'them'.

8.2 De-Idealizing as Far as Necessary and as Little as Possible

If we juxtapose this brief reconstruction of Fricker's genealogy of testimonial justice with the genealogies of Williams and Craig, it becomes apparent that all three genealogies highlight closely intertwined aspects of one system of epistemic division of labour, namely what has recently come to the forefront of epistemology as the institution of *testimony*.[4] The three genealogies complement each other, showing how the concept of knowledge, the virtues of accuracy and sincerity, and the virtue of testimonial justice all have their *raison d'être* in the fact that they are indispensable cornerstones of a system of epistemic division of labour. The three genealogists thereby work towards a comprehensive pragmatic genealogy of our epistemic practices, and in so doing, they erode the boundary between traditional philosophical epistemology that reflects a priori on our most basic epistemic concepts and the more empirically informed investigations into our actual epistemic practices pursued by social epistemology and the sociology of knowledge. Not only can we not understand our epistemic concepts in complete abstraction from contingent facts about human beings and their needs in the kinds of environment they live in, these genealogists maintain; we also cannot understand them in abstraction from the fact that human beings are *social* creatures who interact in and as groups.

It is this last line of thought that Fricker pushes further than Craig and Williams. She shows how the dynamic model of the system of epistemic division of labour can be further refined through de-idealization, i.e. by bringing the model one step closer to the complexities of our actual situation. In particular, the need for the virtue of testimonial justice only comes into view once we de-idealize our model far enough for our epistemic agents to be socially situated. As Fricker notes,

---

[4] See, *inter alia*, Kusch (2004), Gelfert (2014), and Moran (2018). Kusch (2009b), Fricker (1998, 2012, 2016a), Reynolds (2017), and Hannon (2019) develop the genealogical approach to the institution of testimony even further.

'a philosophical framework that prescinds from matters of social identity and power could never give an account' (2007, 177) of that virtue. Much as the mature Nietzsche of the *Genealogy* finds he can improve on his predecessors by replacing their socially homogeneous models with a socially heterogeneous one (*GM*, I, §2), Fricker finds she can add to the genealogies of Craig and Williams by uncovering the practical origins of testimonial justice in social categorization and the prejudicial pressures it brings.

But why stop there with the de-idealization? Is the general methodological lesson of Fricker's advancement over Craig and Williams not that we should de-idealize our pragmatic genealogical model as far as possible, to approximate the intricacies of our social and political reality?

It would be a mistake to conclude that we should always try to de-idealize our pragmatic genealogical model as far as possible. This would be to surrender part of the point of telling a genealogy in the first place; quite certainly, it would be to surrender part of the point of Fricker's genealogy, for it is essential to Fricker's purpose in telling her genealogy that the model be *as generic and idealized as possible*. Having spent a large part of her book showing why cultivating the virtue of testimonial justice would be a good thing for us given the prejudiced climate in which we live, she turns to pragmatic genealogy to situate this demand in a broader philosophical and historical space, relating it to other human needs and purposes and gaining a sense of the extent to which the need for testimonial justice depends on contingencies of our history. Her declared aim is to 'reveal how far the virtue of testimonial justice is a fundamental epistemic virtue—that is, an epistemic virtue serving a purpose which transcends history in that it arises out of an epistemic need that is present in any human society' (2007, 108). This is in line with her characterization of the genealogical method in an earlier piece, where she notes that 'a good genealogical explanation . . . helps us understand to what extent features of our actual practice are necessary, and to what extent they are contingent' (1998, 165). The species of necessity at stake in *Epistemic Injustice*, Fricker tells us (2008, 48; 2010b, 65–6; forthcoming, 9), is the *practically* necessary—not the metaphysically necessary or the 'humanly necessary' (Strawson 1961) that arises out of 'human emotional nature' (Fricker forthcoming, 9). Fricker's genealogy appeals only to 'basic survival needs plus some further social pressures that grow directly out of them' (Fricker forthcoming, 9). This strengthens the genealogy by rendering it less vulnerable to empirical refutation.

Yet it also serves her aim of showing that testimonial justice is not just some twenty-first-century fetish, but is practically necessary given needs that humans have anyway. If the problem to which the virtue of testimonial justice forms a solution were as socio-historically local as that virtue itself—in particular, if testimonial injustice appeared as a problem only to someone who already valued testimonial justice as a virtue—then this practical vindication of the virtue would be too internal to have much force, much as a justification of liberalism that

rested essentially on what people need according to a liberal conception of the person would amount to little more than liberalism patting itself on the back.[5] By contrast, Fricker's genealogy aims to show that the virtue can be practically vindicated in more external terms, as serving 'a purpose which transcends history in that it arises out of an epistemic need that is present in any human society' (2007, 108).

Given this aim, it is clear that we should de-idealize as much as we need to, but as little as we can. For it is in virtue of its abstraction from our concrete situation that the state of nature can lay claim to representing not just a predicament we happen to face given the particulars of our current situation, but one we are bound to face given some of the most general features of the human situation.

8.3 Pairing Genealogical Explanation with a Theory of Error

In Fricker's genealogy, the claim that the virtue of testimonial justice is practically necessary sits alongside the claim that it has in reality largely failed to be realized. In this respect, her genealogy is importantly different from the genealogies we considered in previous chapters in that it presents the virtue of testimonial justice as something that we *should* adopt, but that we have in large part *failed* to adopt. 'History', Fricker wryly remarks, 'has not yet done very much for the virtue of testimonial justice' (2007, 118). Instead of starting from the ubiquity of a conceptual practice and explaining that ubiquity in terms of the importance of the needs to which it answers—as Craig does, for example—Fricker does the reverse: going back to the most general practical origins of our epistemic concepts and virtues, she traces out the development that these would have needed to undergo rather than the development they have in fact undergone, suggesting that testimonial justice is 'something that we can and should aim for in practice' even if we are 'able to achieve it only rather patchily across different sorts of prejudice' (2007, 98–9).

There might seem to be a tension involved in maintaining both that the virtue of testimonial justice is practically necessary given universal and basic human needs and that it has in fact only patchily been achieved. If testimonial justice is practically necessary in the fundamental way Fricker suggests, why has it not materialized more consistently? Its absence might be thought to cast doubt on its alleged practical necessity, inviting much the same concern as that which Émile Faguet expressed when confronted with Rousseau's assertion that man is born free and yet is everywhere in chains: 'it would be equally reasonable to say that sheep are born carnivorous, and everywhere nibble grass' (1891, 41).[6]

---

[5] See Williams (2005c, 94; 2005i, 8).   [6] Quoted in Berlin (1997, 519).

To relieve this tension, Fricker needs to buttress her genealogy with a *theory of error*, i.e. an explanation of why testimonial justice failed to become ubiquitous despite its alleged practical importance. This is an obligation not incurred by genealogies that start out from the ubiquity of a conceptual practice and seek to explain it by appealing to the practical need for such a practice together with the assumption that this need registered in the actual development of the practice. But Fricker takes the widespread *lack* of testimonial justice—and the corresponding near-ubiquity of testimonial injustice—as her starting point. She argues from the (genealogically derived) practical need for testimonial justice, together with the assumption that this need failed to register sufficiently, to the conclusion that we should do more to cultivate and promulgate testimonial justice. But why did this need fail to register sufficiently? Why has history 'not yet done very much for the virtue of testimonial justice' (2007, 118)?

Though she does not present it under this description, Fricker in fact provides material for a theory of error. First, she highlights the covert and protean nature of the prejudices that testimonial justice aims to eradicate: the virtue of testimonial justice is 'bound to be hard to achieve', she notes, 'owing to the psychologically stealthy and historically dynamic nature of prejudice' (2007, 98). Second, what renders the virtue of testimonial justice especially hard to achieve is its dependence on the socio-historical availability of the relevant forms of critical awareness. Gender prejudice, in Fricker's example, can only be reliably neutralized by someone who possesses a critical consciousness of gender dynamics. 'There are circumstances', Fricker remarks, 'under which the virtue *cannot* be achieved, for it is an ethically significant feature of this virtue that it displays a special sort of cultural-historical contingency' (2007, 99). The achievability of the virtue is contingent on the availability in one's society of critical concepts and insights that put one in a position to *know better*. The contingency is ethically relevant, according to Fricker, because one is only *culpably* at fault in failing to exhibit testimonial justice if the relevant form of critical consciousness is available (2007, 100).[7] If testimonial justice failed to materialize as widely as the virtue of respect for property or the concept of knowledge, then, this is because prejudices often work surreptitiously, assuming ever-shifting forms, and thriving unhindered as long as critical awareness of them is lacking. Moreover, the power of prejudice 'is markedly increased in the transition from the State of Nature to historical society' (2007, 120), because with greater social and political complexity comes new fuel for prejudice. And third, our reliance on prejudices in assessing testimony is partially explained by the fact that in the absence of a great deal of background information about a given speaker, we inevitably need to rely on heuristics, including notably stereotypes, in our assessments of credibility. This renders us

---

[7] But see Fricker (2016a), where, mirroring Williams's account of moral responsibility, she allows for a form of *epistemic agent-regret* even where one is not epistemically at fault.

perpetually susceptible to mistaking a prejudiced stereotype for an empirically reliable stereotype (2007, 30).[8]

Owing to this tripartite theory of error, Fricker's genealogy can still be said to be genuinely explanatory of the actual state of testimonial justice: it accounts for the merely partial emergence of testimonial justice by highlighting not only pressures driving its emergence, but also pressures hindering it.

8.4 Making Ameliorative Use of Pragmatic Genealogy

Although it can be seen to have an explanatory aspect once paired with a theory of error, it is the ameliorative aspect of Fricker's genealogy that makes it stand out: it is more indicatory of what we should strive for than vindicatory of what we have. The pragmatic genealogies we have considered so far have focused on making sense of why we have the concepts we have rather than on (re-)engineering better concepts—in Carnapian (1950, 6) terms, they have been about why we think in terms of *fish* in the first place rather than about getting us into the scientifically fruitful habit of thinking in terms of *piscis*. This changes with Fricker, who thereby brings the pragmatic genealogical tradition into the vicinity of Sally Haslanger's work. Haslanger explicitly discusses the possibility of using genealogy for ameliorative purposes. She even gives a characterization of ameliorative genealogy that is broad enough to encompass Fricker's genealogy. An '*ameliorative genealogy*', Haslanger writes, 'undertakes to evaluate the point of having a concept or structure of concepts (along with related practices) and proposes improved resources to fulfill them' (2012a, 372).

Yet a gap between Haslanger's and Fricker's way of combining pragmatic and genealogical approaches remains. For one thing, Haslanger's genealogical approach is modelled on the rational reconstructions of Carnapian explication (2012b, 367n1), which means that while it combines an interest in how to define a given concept with a pragmatic concern for how well candidate definitions will serve our purposes, it remains primarily in the business of providing answers to Socratic 'What is X?' questions.[9] More significantly, the genealogies Haslanger appears to have in mind are firmly historiographical rather than state-of-nature based—closer

---

[8] An important complication, which I cannot pursue further here, is that even empirically reliable stereotypes may present problems of their own once we consider *why* they are empirically reliable, and what harm even a reliable stereotype may cause. See, e.g., Basu (2019).

[9] Haslanger distinguishes between a *conceptual*, a *descriptive*, and an *ameliorative* approach to 'What is X?' questions (2012b, 367). The first looks at our *intuitions* about a concept and what falls under it; the second looks at what *kinds* (if any) a concept tracks; the third inquires into the *point* of the concept. On Haslanger's view, all three approaches can be transposed into a genealogical key (2012b, 371). See Brun (2016, 2020), Dutilh Novaes (2018), and Mühlebach (2016) for discussions of Haslanger's view and its Carnapian roots.

to Foucault and Hacking than to Fricker and Hume.[10] On Haslanger's conception, ameliorative genealogy seeks to identify the point of a concept by considering its workings in a series of different but real historical circumstances. Despite these differences, however, both Fricker and Haslanger propose to use genealogy in order to improve our conceptual resources so as to better fulfil our critically examined needs and purposes.

The claim that Fricker's *genealogy* is ameliorative should not be confused with the claim, which is also true, that its *object*, the virtue of testimonial justice, is ameliorative or corrective. The virtue serves to counter pressures towards testimonial injustice, and this has implications for Fricker's approach: the point of testimonial justice only reveals itself against the background of an understanding of *what can go wrong* in the system of epistemic division of labour. It is by first grasping what the pressures towards failure are that one can come to see 'what counter-pressures the structure needs persistently to exert in order to stave off collapse into the negative' (Fricker 2015, 73); the 'functional forms of things need to be seen as successfully staving off or coping with endemic problems and difficulties' (Fricker 2017, 57). This is what Fricker elsewhere calls a 'failure-first approach' (2015, 73). She writes that 'a philosopher who only aimed to understand and represent epistemic practices in their most functional forms, perhaps even in some notionally ideal form, would still need to do so by looking first at what potential collapses into dysfunctionality are being perpetually staved off, and by what mechanisms' (2017, 57). The idea here is presumably not that we should identify the function of something by looking first at what—in relation to this function—would count as instances of malfunction, since this would presuppose what it is trying to show. Rather, what is being recommended is an approach akin to the one characterized at the beginning of this book: in trying to understand a conceptual practice, we begin by looking for *problems* that it might form a response to, and let our understanding of its function grow out of our understanding of these problems. But although testimonial justice is a corrective virtue, it is not the corrective nature of its object that makes the genealogy itself corrective; accuracy and sincerity both centrally involve resisting various pressures towards laziness, wishful thinking, and deception, and yet Williams's genealogy is for the most part vindicatory of what we already have rather than indicatory of what we should strive for.

The ameliorative argument at the heart of Fricker's genealogy is the following: (i) if the system of epistemic division of labour that we are all committed to already in virtue of our most basic human needs is to *function well*, it needs to be free of the distorting influence of prejudice; (ii) the practices we in fact have *fall far short* of achieving this even approximately; (iii) a *salient remedy* to this problem,

---

[10] See Haslanger (2012a, 19, 369–79) and Hacking (1999, 10–14).

which we have a *pro tanto* reason to adopt already insofar as we have an interest in participating in a well-functioning epistemic division of labour, is to cultivate the virtue of testimonial justice to counteract the distorting influence of prejudice.

What this argument suggests is that a system of epistemic division of labour that was maximally efficient would also need to be maximally resistant to confounding influences, and would therefore see the virtue of testimonial justice at work everywhere and all the time, completely neutralizing the impact of prejudice (though not necessarily the impact of stereotyping). As Fricker demonstrates in the rest of her book, there are a number of reasons to want to avoid testimonial injustice—many of them more ethical than epistemic or prudential in character; but what the genealogical story in particular underlines is that there is a rationale for the virtue of testimonial justice that is baked into our epistemic division of labour already at its most basic level. This is a reason to cultivate the virtue that even people who are otherwise unmoved by the various ethical reasons to avoid testimonial injustice—people who do not care about, or stand to profit from, the harm done to the victims of testimonial injustice—could in principle recognize as a reason for them. The genealogy appeals to an uncontroversial because extremely widely shared element in the motivational set of human agents, namely the interest that all human agents have in getting at truths about their environment, and hence in participating in a well-functioning practice of information pooling. As Fricker puts it,

> it could not be controversial to assume that epistemic subjects considered as such possess in their actual set of motivations some general motivation to truth, and *a fortiori* some motivation to more proximal ends which are in the service of truth (such as neutralizing prejudice in one's habits of trust, for instance)....All may agree that, in general, any epistemic subject will have a reason to get at the truth....Even the most virulent, dyed-in-the-wool sexist version of Herbert Greenleaf, possesses a motivation (to truth) from which there is a sound deliberative route to questioning his spontaneous lack of trust in Marge.   (2007, 102–3)

In short, the genealogy identifies a motive for neutralizing prejudice which even the most unabashedly prejudiced person is likely to share. Being prejudiced has epistemic costs. One loses out on many truths that might be important to one, but also on other epistemic goods such as fruitful questions, suggestions, and objections.

On balance, of course, this genealogically derived insight into the epistemic costs of being prejudiced may not carry all that much weight with the prejudiced person, and I believe Fricker is under no illusion that the fear of missing out on a few truths will convert such a person into a paragon of testimonial justice. This, presumably, is why so much of the book is devoted to explicating the various other

things that are wrong with being prejudiced besides the epistemic loss it entrains. Fricker is aware that to *reduce* the phenomenon of testimonial injustice to a problem of epistemic loss is *already* to be prejudiced. The persuasiveness of Fricker's case for testimonial justice stems from the cumulative force of the various reasons she works through—some weaker but widely shared, some stronger but less widely shared. Most of these reasons have no immediate connection to the genealogical story. The weightiest among them have to do with the various forms of harm incurred by those who are unjustly deprived of the credibility they deserve.

The genealogy nonetheless achieves several things that could not have been achieved simply by examining the harms done by testimonial injustice. One is that it presents testimonial injustice as an *utterly basic* form of epistemic injustice—indeed as 'being, at least genealogically speaking, the most basic of all' (Fricker 2010a, 174). This lends succour to Fricker's claim that 'the relevant category of epistemic injustice under which to gather a range of wrongful exclusions from the discursive practice of inquiry is indeed "testimonial injustice"' (2010a, 176). We shall come back to the question of what something's being basic in a genealogical story really tells us about its position within our actual practices, but for now let us note that Fricker's genealogy serves to 'establish the taxonomical propriety of gathering diverse forms of prejudicial exclusion from discursive participation under the general head of "testimonial injustice"' (2010a, 176).

Second, the genealogy provides us with a powerful integrative model through which to discern and connect the various ways in which testimonial justice does something for us. Function ascriptions are inevitably holistic, which is to say that something has a determinate function not intrinsically and in isolation from everything else, but only as part of a certain natural and social environment.[11] This can make it difficult to identify the various beneficial effects of something in the overwhelmingly complex environment we actually live in. It is even more difficult when we are asking what beneficial effects something we largely lack *would* have. It can therefore be helpful to *construct* a simplified state-of-nature environment that perspicuously displays the various beneficial effects of something—especially if, as in the case of testimonial justice, the function ascription operates with a novel concept, so that a firm conceptual grip on the relevant phenomena cannot be taken for granted. It is fairly obvious that testimonial justice does something for the person who is at the receiving end of it. What is less obvious is that it also does something for the individual who cultivates it and for the flow of epistemic goods in the community as a whole (which in turn tends to

---

[11] See, e.g., Kincaid (2020, 24). See also Hufendiek (2020, 101), who argues on this basis that, especially when functional explanations are offered to promote self-understanding rather than to facilitate precise prediction or manipulation, it makes sense to resort to the rational and interpretive modelling exemplified by state-of-nature models.

benefit its members). Neutralizing their prejudices gives individuals access to truths they would otherwise miss out on, and as the case of Herbert Greenleaf poignantly illustrates, some of these truths may be of great value to these individuals. But even where they are not, the individual's cultivation of the virtue of testimonial justice serves a function from a social point of view, because as the genealogy makes clear, testimonial justice is a key contributor to the success of collaborative inquiry—any prejudice among the members of the epistemic community is a weak spot in the system, a flaw that threatens to hamper the growth or vitiate the quality of the pool of information. The genealogical model helpfully renders these relations between prejudices, individuals, and social needs perspicuous.

A third and related feature of the genealogical model is the one familiar from Chapter 7 on Williams—that it can help us understand why testimonial justice really needs to be a *virtue* rather than a mere disposition. Fricker does not directly address this issue, but it is clear that she takes testimonial justice to be a 'virtue of truth' (Fricker 2007, 6) in something like Williams's sense. Drawing on Linda Zagzebski's (1996) account of virtue, Fricker writes that 'to qualify as possessing the virtue', a person 'would need to have a suitably entrenched general motivation to make unprejudiced credibility judgements' (2007, 93). This, by itself, does not yet tell us whether the motivation is to be instrumental or intrinsic—one might be motivated to neutralize prejudice and be reliably successful at it for purely instrumental and self-regarding reasons, driven solely by a concern to garner information that is useful to oneself. But of course, from such an instrumentalizing stance, only certain prejudices would seem worth neutralizing in the first place, and then only when doing so happened to align with one's own interests on a given occasion. As a result, all the benefits *to others* that a less selective form of testimonial justice might have carried would be lost. One might object that such selective testimonial justice is barely intelligible, since beliefs are not subject to the will in a way that would allow one to *decide*, once one became aware of the possible influence of prejudice on a given occasion, whether or not to bother neutralizing it—to become aware of a looming prejudice is *already* to question one's initial credibility judgement, and there is little room for the thought 'I'm prejudiced here, but I'll discount her testimony anyway'. But it remains the case that a voluntary effort is often involved in first bringing oneself to even consider the possible influence of prejudice on a given occasion. The purely self-regarding practitioner of testimonial justice may simply not bother to make this effort unless driven by personal interest. Critical reflection is work, and there is nothing unintelligible about the person who is content to be comfortably uncritical unless they are jerked out of their complacency by the realization that their own welfare depends on their getting it right.

Since Fricker explicitly builds on Williams's genealogy, it is therefore natural to see her account as buttressed by his genealogical explanation of why the virtues of

truth really need to be *virtues* if they are to fulfil their function. It is only if testimonial justice is valued *for its own sake*, as an expression or realization of truthfulness, that it can motivate people to neutralize prejudice even when they have otherwise no reason to do so. In line with this, Fricker notes that the aim of neutralizing prejudice is a 'proximal' aim which is in turn made sense of as conducive to achieving a further end, namely truth (2007, 99). And as we saw Williams maintain, to value the truth intrinsically is to value the various expressions of what he brings under the heading of truthfulness. Fricker demonstrates that an important expression of truthfulness is the neutralization of one's prejudices. If Williams's genealogy shows that we need to value the truth intrinsically and that this comes down to understanding truthfulness as a virtue, as a disposition worth realizing for its own sake, Fricker's genealogy shows that this includes regarding testimonial justice as a virtue: like accuracy and sincerity, testimonial justice should be regarded as an intrinsically valuable expression or realization of truthfulness.[12]

Fourth, the genealogy shows us the deeper ramifications of testimonial justice and injustice, and thereby reveals the cultivation of testimonial justice to be connected, in potentially unsuspected ways, to other things that we value, such as free speech, equality, and non-domination. The neutralization of prejudice is shown to be a fundamental functional demand on any system of epistemic division of labour, and once we grasp the fundamentality of testimonial justice to human life—its basic role at the very root of a system we all have an interest in being part of—we come to see the many ways in which it ties in with other things that we value. This means that testimonial justice also derives value from other goods whose realization partly depends on the realization of testimonial justice. For example, Fricker (2013) argues that once we appreciate the fundamental practical significance of testimonial justice, we can come to recognize its connection with the capacity to make one's voice heard and to achieve a culture of free speech; we can likewise come to recognize its connection with the capacity to contest domination and to achieve political freedom and equality (2015). Reconstructing the exact manner in which she draws these connections would take us too far afield, but the important methodological point is that the normative conclusion that we should cultivate testimonial justice is supported not just by the basic epistemic needs that demand an epistemic division of labour, but also by more local needs, such as the needs for free speech, equality, and non-domination. The genealogical model helps us to see the internal relations between the pursuit of these goods and the cultivation of testimonial justice.

Finally, perhaps the most fundamental motivation for tracing the pragmatic genealogy of the virtue of testimonial justice and connecting it with its Williamsian

---

[12] One could also imagine classifying testimonial justice as an important but easily overlooked elaboration of Williamsian 'Accuracy' rather than as a separate virtue of truth.

and Craigean prequels is that it vindicates social and feminist epistemology against the more classical forms of epistemology that have tended to ignore matters of social identity, power, and prejudice. Fricker's genealogy vividly reveals questions of social identity, power, and prejudice to be not just peripheral or epiphenomenal to epistemology, but to cut to its very root: even the most basic and primitive form of the concept of knowledge imaginable already requires one to negotiate the distorting influence of social categorization, and the proto-virtue of Testimonial Justice is revealed to be (along with Accuracy and Sincerity) among the absolutely basic epistemic virtues. In light of this genealogical understanding of our system of epistemic division of labour, Fricker concludes, we come to see that

> what the recipient of a prejudicial credibility deficit is excluded from is the single practice that dramatizes the origin of what it is to be a knower at all. Testimonial injustice denies one access to what originally furnishes status as a knower. No wonder, then, that even relatively inconsequential testimonial injustices can carry a symbolic weight to the effect that the speaker is less than a full epistemic subject: the injustice sends the message that they are not fit for participation in the practice that originally generates the very idea of a knower.   (2007, 145)

Of course, just *how much* weight this point carries will depend on how central to our epistemic practices the institution of testimony actually is, and as we saw in previous chapters, the mere fact that testimony is at least usefully *modelled as* originally central does not by itself suffice to show that it *still* is. For her part, Fricker follows Craig in thinking that the prototypical epistemic practices of the state-of-nature model also form the core of our actual practices: 'When someone is excluded from the relations of epistemic trust that are at work in a co-operative practice of pooling information, they are wrongfully excluded from participation in the practice that defines the core of the very concept of knowledge' (Fricker 2007, 145). Yet if one thinks of pragmatic genealogy as presenting us with a prototype and its actual elaboration rather than with the core of our actual practice and its additional layers, this overstates the case. It is one thing to show that testimony figures centrally at the base of a plausible model of why we came by various elements of our epistemic practices; it is quite another to show that it also figures centrally in the epistemic practices we actually have. Indeed, as Kusch (2009a, 173) highlights, the originally close connection between the concept of knowledge and testimony is *weakened* in the course of Craig's story, so that his genealogy itself suggests that testimony would in fact end up being less central than it initially was. Once Craig factors in the possibility of group actions, it becomes clear that there are contexts in which inquirers care only about whether *someone* in the group proto-knows (think of mushroom picking, where only one person needs to know which mushrooms are edible). Inquirers 'cease to care

whether the needed information is accessible to them as individuals; they are satisfied if it is accessible to someone in the group', and 'speak of protoknowledge even outside the context of testimony' (Kusch 2011, 9). As a result, the necessary connection between being a proto-knower and giving testimony is lost. Even on this dynamic model interpretation of Craig's genealogy, however, we can still conclude with Fricker that to be excluded from the system of testimony is to be excluded from the practice that 'dramatizes the origin of what it is to be a knower at all' (2007, 145).

In sum, there is plenty in Fricker's genealogy—not to speak of the non-genealogical parts of her arguments—to support the conclusion that we should do more for testimonial justice. This ameliorative conclusion is supported not just by the genealogical insight into its foundational role in the practice of information pooling, but also by the various concerns it ties in with in virtue of that foundational role. If testimonial justice is crucial to the practice of pooling information, and this practice is in turn crucial to other goods, then testimonial justice is crucial also to these other goods. It is not least by highlighting this dependence structure and channelling the normative force of seemingly unrelated concerns towards the increased cultivation of testimonial justice that the genealogy acquires its persuasive power.

From a methodological point of view, the genealogy vividly shows how much additional insight can be drawn out of a genealogical model merely by factoring in social categorization and breaking up the homogeneity of the state-of-nature community. But the genealogy's more ground-breaking contribution to the pragmatic genealogical tradition is that it shows how genealogical reverse-engineering can feed into forward-looking conceptual engineering by encouraging us to cultivate something we at least partly lack. Fricker uses the genealogical method to dig up the important function that testimonial justice *would* discharge within the system of epistemic division of labour *if* that system developed in more effective ways; she diagnoses and explains the widespread failure of the system to develop in that way and the corresponding lack of testimonial justice it entrained; and she then uses this idealized, counterfactual pragmatic genealogy as a guide to the amelioration of our practices.[13]

With Fricker's strategy, there are thus three ways in which conceptual reverse-engineering can feed into conceptual engineering. First, in revealing *what the point* of a conceptual practice is and *which instances* of the practice paradigmatically serve that point, one also marks out the instances in which it is *pointless* or *overreaches* itself. With a critical eye for pointfulness comes a critical eye for pointlessness, and amelioration can take the form of cultivating the pointful at the expense of the pointless. Second, in revealing the importance of the *context of*

---

[13] This 'reverse-engineering as a guide to engineering' strategy is explicitly recommended by Thomasson (2020).

application in giving point to a conceptual practice, one calls its *tel quel* application beyond that context into question, indicating that if the practice is to have a point in the new context, the practice must be re-engineered and adapted in such a way as to reproduce its pointfulness in the new context. The question then is what new elaboration a practice demands in order to serve the same point in a different context—a question pursued by Damian Cueni (2020), for example, when he asks how we should re-engineer our ideas about what counts as legitimate rule *within* the nation state for use *beyond* the nation state. By reverse-engineering what those ideas do for us in the domestic source context, Cueni argues, we gain a guiding sense of what forms of functionality we should seek to safeguard and recreate as we extend and adapt these ideas of legitimacy to the international target context. Third, as Fricker illustrates, one can identify the practical pressures at the root of a conceptual practice and examine how the practice *would* have developed *if those practical pressures had been given free rein*, and then use that counterfactual genealogy to guide efforts to ameliorate our practice as it actually developed. In each case, to reverse-engineer and go backwards to the practical origins of the practice can be a way of preparing and giving direction to a forward-looking amelioration of the practice. Then, as Nietzsche put it, one goes 'backwards as everyone goes backwards who wants to take a big jump' (*BGE*, §280).

Of course, this prospective exploitation of retrospective genealogical reflection also presents difficulties of its own. Nowhere are these difficulties more vividly illustrated than in the project ostensibly contemplated on occasion by Nietzsche himself, of engineering an entire new set of values *ex nihilo*.[14] We do not fashion our concepts out of nothing, but are bound to draw on the concepts we find ourselves with—both in fashioning new concepts and in evaluating what these concepts should look like. And even if genealogical reflection can yield a fairly specific functional outline of the kinds of ideas we have reason to adopt, there remains a question as to whether the mere recognition that one has good reason to adopt and internalize a particular virtue or value suffices actually to do so. This is not just a matter of the limitations of the individual's capacity to change what is essentially shared and social. It is also matter of the sense-making constraint we saw Williams highlight—the constraint that for something genuinely to become a virtue or value one lives by, the virtue or value has to *make sense from the inside*, and whether or not something makes sense to us in this way is not a matter of our will, but rather 'comes as a discovery' (Williams 2002, 261–2). If the recommended virtue or value is to make sense, it needs to be embedded in a structure that allows it to engage our emotions and us to articulate why it matters.

Yet despite these obstacles, Fricker's project manages to strike a balance between her ameliorative ambition and the recognition that we must work with

---

[14] See Williams (2000, 2006i) and Queloz (forthcoming-a) for critical discussions of that project.

the material we find. She does not purport to fashion a virtue *de novo*, but rather aims to strengthen and extend the reach of a virtue that we—as a community, and only *in nuce*—already have. She offers us a vindicatory explanation of why we came to have this virtue at all, and, by relating it to various goods that we already value, shows us why it is worth cultivating more broadly. This makes her ameliorative project look a lot more feasible than the more extreme Nietzschean versions of genealogically informed value engineering. The sense-making constraint is no bar to her enterprise, since, in the form of values such as truth, justice, and freedom, we already possess the material necessary to make sense of testimonial justice as a virtue. By uncovering the ramifications of testimonial justice and the way it contributes to the realization of other goods, the genealogy *shows* us that it makes sense to us as a virtue. And by tracing the virtue's practical origins to basic needs for information and cooperation, the genealogy shows us that it *makes good naturalistic sense* that testimonial justice should make sense to us as a virtue. Like the other vindicatory genealogies we encountered, Fricker's genealogy works to strengthen one's confidence in its object in several ways at once. It presents us with a series of reasons—some widely shared, some less so—to cultivate the virtue of testimonial justice; and it also works to strengthen our confidence in testimonial justice in a subtler way, namely by showing us that this virtue, for all the neologistic and technical character of Fricker's label for it, is really nothing mysterious or new-fangled: it fits seamlessly into a naturalistic picture of humans as cooperative inquirers, and the practical pressures on it to arise become visible as soon as one reflects, even in the most general terms, on the demands of cooperative inquiry in a socially heterogeneous community. Moreover, there is a performative power in the telling of such a genealogical story itself, a power that goes beyond the argumentative force of the reasons to cultivate testimonial justice which the story brings to light. The story itself increases our *receptiveness* to these reasons by dispelling our suspicions and making us comfortable with testimonial justice, assuring us that testimonial justice has a natural and rightful place in human affairs.

But there is also, finally, a performative aspect to Fricker's genealogy which it does not share with any of the other genealogies we discussed, namely that whoever hears or reads the genealogy thereby acquires a new thinking tool: the *concept* of the virtue of testimonial justice as opposed to the virtue itself. Acquiring the concept of testimonial justice renders us sensitive both to the possibility of consciously cultivating it and to its actual presence or absence in ourselves and others. Moreover, having a label by which to express and refer to that new concept further empowers us to engage with others in discerning and encouraging the cultivation of testimonial justice. Reflecting on her book in a later piece, Fricker explicitly notes that one of her chief concerns had been to forge an 'on-the-ground tool of critical understanding that was called for in everyday lived experiences of

injustice' (2017, 56). The already vast and still rapidly growing literature on testimonial justice and its cognates suggests that she has succeeded.[15]

For the philosophical methodologist, this holds two broader lessons. One is that the genealogical method can be used not only to strengthen or weaken our confidence in existing concepts, but also to generate and instil new concepts. The other is that an ameliorative genealogy that advocates a certain improvement in our conceptual practices also has a more immediate ameliorative effect: by telling it, one equips others to pursue the amelioration it recommends.

With this last case study in pragmatic genealogy, we have come all the way from Hume to the present day. It is time to switch back to a more purely systematic perspective to address various objections and to examine in more detail wherein the normative significance of pragmatic genealogies lies.

---

[15] See, e.g., Anderson (2012); Berenstain (2016); Dotson (2011); Kidd and Carel (2017); Origgi (2012); Pohlhaus Jr (2014); Thomas (2018); Wanderer (2012); see also the dedicated issue of *Episteme* (Alcoff 2010; Coady 2010; Cullison 2010; Fricker 2010a; Goldberg 2010; Hookway 2010) and the essays in *The Routledge Handbook of Epistemic Injustice* (Kidd, Medina and Pohlhaus Jr 2017).

9

The Normative Significance of Pragmatic Genealogies

Genealogical explanations are widely thought to be normatively inert: they may advance our understanding of philosophically puzzling concepts, but they do not directly vindicate or debunk them—they are only preparatory, but not constitutive of critique.[1] Accordingly, attempts to use genealogy to affect the space of reasons are frequently dismissed as committing the 'genetic fallacy'.[2] Against this standard view, I have suggested that pragmatic genealogies can be carefully constructed to showcase the respects in which certain conceptual practices serve our needs, and that in virtue of this, they can be normatively significant and affect the space of reasons. But in what sense exactly? To carve out the contours of this claim, we can chisel it with four increasingly refined objections: (i) normatively ambitious genealogies commit the genetic fallacy; (ii) if they do not commit the genetic fallacy, they founder on failures of continuity in the conditions from which they draw their normative import; (iii) if they do not founder on such continuity failures, this must be because they are based on anthropologically universal needs, which severely restricts their explanatory scope; and (iv) even this does not provide a solid basis, for need ascriptions are contestable (as illustrated by the economists' jibe that a 'need' refers to something you want, but are not prepared to pay for).

As we work through these four objections to substantiate the thesis that pragmatic genealogies can affect the space of reasons, we will also have occasion to tie up some loose ends. In particular, the notions of *pointfulness*, *needs*, and the *conception of the agent* with which pragmatic genealogies operate will be clarified. It will also emerge that pragmatic genealogies can model even entirely socio-historically local problems, and help us distinguish between different ways in which our ideas seem to be necessary for us.

---

[1] See Dutilh Novaes (2015, 100–1) and Koopman (2009).

[2] This objection was raised against Williams's genealogy by Koopman (2013, 20). For further discussions that see the genetic fallacy as clouding the prospects of normatively ambitious genealogical explanations, see Fraser (1981), Glock (2008a; 2008b, 101), Goudge (1961), Hanson (1967), Kaplan (2002, 13), and Kim (1990).

9.1 Genetic Fallacies and the Ways Around Them

The first objection to the claim that pragmatic genealogies can affect the space of reasons, which is not yet specific to pragmatic genealogies and calls for an answer in more general terms, maintains that attempts to derive reasons from claims about the genesis of something commit the genetic fallacy. This alleged fallacy is well known; what is less so is that when Morris Cohen and Ernest Nagel introduced the term in 1934, they identified not one, but *two* distinct genetic fallacies. And as Amia Srinivasan (2019, §1n3) notes, it is the first and lesser-known of the two that poses the greater *prima facie* threat to vindicatory genealogies.

This first form of the genetic fallacy 'takes a logical for a temporal order' (Cohen and Nagel 1934, 388): it fallaciously infers from the fact that something is *logically prior* that it must be *temporally prior* as well. To condemn this form of reasoning as a fallacy, Cohen and Nagel think, is to condemn

all attempts current in the eighteenth century, and still widely popular, to reconstruct the history of mankind prior to any reliable records, on the basis of nothing but speculations as to what must have been. The theories as to the origin of language or religion, or the original social contract by which government was instituted, which were based on empirically unsupported assumptions as to what 'the first' or 'primitive' man *must* have done are all historically untenable. It is clearly a logical error or fallacy to assume that actual history can be so constructed or discovered. Not much different, however, are those speculative *a priori* histories which under the name of social evolution attempt to deduce the stages which all human institutions must go through and therefore actually have gone through. In all of these attempts to trace the history of the family, industry, the state, and the like, the earlier stages are assumed to have been simpler, and the later stages more complex. Such attempts appeal to us because we can understand the present complex institutions better if we see them built up out of simpler elements. But it is an inexcusable error to identify the temporal order in which events have actually occurred with the logical order in which elements may be put together to constitute existing institutions. Actual recorded history shows growth in simplicity as well as in complexity.

(Cohen and Nagel 1934, 388–9)

Even leaving aside the respects in which Cohen and Nagel are representative of the verificationists of the 1930s, who have their own particular reasons to be unsympathetic to unverifiable speculations about the past (1934, 207–11), their objection looks to be fatal to state-of-nature-based genealogies *if these are interpreted as conjectural histories*. But if they are interpreted as *dynamic models* whose order is in the first instance meant to reflect the order of optimal intelligibility rather than

the temporal order of actual development, the objection loses its force. Indeed, it is the point of pragmatic genealogies understood as dynamic models to exploit the fact that 'we can understand the present complex institutions better if we see them built up out of simpler elements' *without* committing ourselves to groundless speculations about actual history. As became clear in our discussion of Craig and Williams in Chapters 6 and 7—though it is already true of Hume as interpreted in Chapter 4—these vindicatory pragmatic genealogies start out from states of affairs that, so far from being alleged to have obtained at some early stage of history, are not alleged to have obtained at all—indeed, they are alleged *not* to have obtained, because they involve unrealistic or unstable idealizations. Instead of being a threat to pragmatic genealogies, therefore, Cohen and Nagel's first form of the genetic fallacy precisely brings out the advantage that, on the dynamic-model interpretation, pragmatic genealogies possess over conjectural histories. By *constructing* a progression from the simple to the complex in a model instead of looking to history to offer such a progression in ready-made form, pragmatic genealogies elegantly sidestep this genetic fallacy, reaping the benefits of Enlightenment-style origin stories without the costs.

As for the second form of the genetic fallacy that Cohen and Nagel introduce, which is the currently prevalent understanding of it, it boils down to the error of treating items in the *context of formation* of conceptual practices as if they belonged to the *context of justification* when in fact they do not.[3] We can acknowledge that there is such an error without committing ourselves to the much stronger claim that nothing can be inferred about the justification of something from facts about its origins. Items in the context of formation *can* form part of the context of justification, but they can properly do so only if there is a *connection* between some aspect of the context of formation and the justification of the item in question. What is fallacious is not the inference from origins to justification *per se*, but the inference from *irrelevant* information about origins to justification. And of course, *whether* some piece of information is relevant to the justification of a given conclusion is often precisely what is at issue.[4] What normatively ambitious genealogical explanations depend on, then, is that there be a connection rendering some aspect of the context of formation

---

[3] This Reichenbachian characterization is adapted from Salmon (1973, 11). Alternative characterizations of the fallacy maintain that it consists in judging the truth of an assertion on the basis of its source rather than by the evidence available for it (Kaplan 2002, 13), or in conflating temporal origin with logical nature (Koopman 2013, 20).

[4] The rise of the genetic fallacy charge in the 1930s and 40s may have been a response to the genetic reasoning involved in anti-Semitic discreditations of 'Jewish science' (Giere 1999, 14). But the distinctions animating the charge—between genesis and validity, explanation and justification, causes and reasons—were already in high demand during the 'psychologism' debates that raged from the 1880s to the 1920s, when philosophers needed to demarcate their discipline from the emerging field of psychology (Kusch 1995).

relevant to the context of justification.[5] Such a connection can be forged from two directions. Either the space of reasons is itself such that it locally encompasses certain formation processes, because something is justified or claims authority for itself *in terms of* its formation; or the formation processes are such that knowledge of them can contribute to the vindication or subversion of their product even when that product was not already justified in terms of its formation. Either we have *genetically justified practices*, or we have *genealogies yielding justifications*.

Consider genetically justified practices first. Claims about the formation processes of conceptual practices may affect the space of reasons because these claims concern practices whose *authority is itself a function of their formation*.[6] That is to say, formation processes are part of the truth conditions of the propositions from which the practice derives its authority. We may call practices that understand themselves or claim authority for themselves in terms that knowledge of their formation can undermine *genetically justified practices*. Examples of such genetically justified practices abound in politics and law, where it is common for practices to derive their authority from the procedure by which they were formed. There are also many rituals and traditions that justify their continuation by reference to their authoritative origins—things are done a certain way because some respected originator did them that way. Religious practices in particular tend to revolve around widescreen representations of their own origins from which they derive their self-understanding and authority.

In such cases, genealogical explanations can impinge on the space of reasons because the rational articulation of the practices in question itself refers to their formation. It is in virtue of displaying such a justificatory connection to their own formation that certain practices will be susceptible to vindication and subversion by genealogical explanations. The structure of such vindications and subversions will then be as follows:

Vindicatory Genealogy of a Genetically Justified Practice:

Conceptual practice *P* claims authority for itself in terms of a representation R_{FP} of its own formation process *FP*.

Inquiry into how *FP* might have given rise to *P* suggests that R_{FP} is true.

Therefore, the authority of *P* is to that extent vindicated.

---

[5] Pashman (1970) argues that relevance depends on there being a causal link between the context of formation and the context of justification. But in many cases, no such philosophically neutral ways of determining relevance will be available: the Archimedean standpoint is lacking (Crouch 1993; Srinivasan 2015).

[6] Williams (2014g, 410) and Gutting (2005, 50).

Subversive Genealogy of a Genetically Justified Practice:

Conceptual practice P claims authority for itself in terms of a representation R_{FP} of its own formation process FP.

Inquiry into how FP might have given rise to P suggests that R_{FP} is false.

Therefore, the authority of P is to that extent subverted.

Genealogy can thus sidestep the genetic fallacy by exploiting the fact that the target phenomenon understands itself and claims authority for itself in terms that render it vulnerable to genealogical inquiry: genealogical inquiry can sap the authority of beliefs or ideas *insofar* as these demand authority for themselves in terms that are incompatible with the kinds of origins that genealogical inquiry shows them to have. Thomas Nagel (1997, 2009a) renders the authority of liberalism vulnerable to genealogical subversion, for example, when he represents liberalism as the rationally inevitable product of a historical process of attunement to universal reasons.[7] Wherever there is such a veneer of inevitability, genealogy can peel it away by generating a sense of alternative possibilities: people *can* live and reason differently, because they *have* lived and reasoned differently.[8]

However, there are other ways in which genealogy can impinge on the space of reasons which we miss if we focus only on how genealogical revelations of contingency can undermine claims to inevitability or necessity. These are not a matter of how practices claim authority for themselves, but of how and why they originated. This is where we turn from genetically justified practices to *genealogies yielding justifications*.

In this second way of connecting origin and justification, it is not the justificatory structures but the formation processes themselves that are such that knowledge of them can contribute to the vindication or subversion of practices, or simply exhibit them as rationally contingent.[9] Let us say that a practice P is *rationally contingent* to the extent that the considerations contributing to the best explanation of why a group G engages in P fail to provide *reasons* to prefer P over possible rivals to P, where *possible rivals* to P are unrealized alternatives to P competing for the place in our lives occupied by P, and notably include the abandonment of P. We can then distinguish three ways in which insights into the formation process of a practice can bear on our understanding of it:

---

[7] See Williams (2014g, 410).

[8] Even when a practice appears inevitable or necessary *without* its authority depending on its being taken to be so, genealogical inquiry can prompt a critical reevaluation of authority. Generating a sense of alternatives will then not itself constitute a subversion of authority. But, as Elizabeth Anderson puts it, it can convert dogmas into tools that we can choose to use or not (2001, 22). It can turn dogmatic acceptance into critical assessment.

[9] As Srinivasan (2015, 2019, manuscript) argues, *which* formation processes count as vindicatory, subversive, or neutral is a question that cannot ultimately be answered without drawing on some first-order commitments concerning what one in fact takes to be true or valuable.

Vindicatory Genealogy:

Group G engages in conceptual practice P.

The best explanation for why G engages in P is that P is the result of formation process FP.

FP is vindicatory, i.e. it offers reasons to prefer P over possible rivals, including the abandonment of P.

Therefore, the continuation of P is to that extent justified.[10]

Non-Vindicatory Genealogy:

Group G engages in conceptual practice P.

The best explanation for why G engages in P is that P is the result of formation process FP.

FP is not vindicatory, i.e. it fails to yield reasons to prefer P over possible rivals, including the abandonment of P.

Therefore, P is to that extent rationally contingent.

Subversive Genealogy:

Group G engages in conceptual practice P.

The best explanation for why G engages in P is that P is the result of formation process FP.

FP is incriminating, i.e. it offers reasons against the continuation of P.

Therefore, the abandonment of P is to that extent justified.

There are two reasons for reconstructing the genealogies at issue here as forms of *abductive* reasoning, i.e. reasoning to the best explanation. The first is that it enables us to offer genealogies even where we have no knowledge of how a given practice actually came about (knowledge that a sound *deductive* argument would require). The second reason is that the abductive reconstruction licenses a form of self-referential reasoning (Klement 2002, 392): the existence of the practices can be used as evidence for the existence of the formation processes imbuing them with authority, which is important because the existence of these formation processes is often just what is at issue. At the end of Tolstoy's *Anna Karenina*, for example, Levine comes to see the existence of his own moral values as evidence for the truth of divine revelation, a formation process whose existence in turn vindicates those values (Tolstoy 2014, VIII, chs. 12–13). Another example is Descartes's argument for the existence of God in the *Third Meditation*: among the contents of his consciousness, Descartes finds, is the concept of unified

---

[10] Here I am generalizing to conceptual practices a pattern of genetic reasoning spelled out by Klement (2002, 390).

perfection, i.e. the concept of God; this concept of unified perfection could not have come from something less than perfect; since he himself is imperfect, the best explanation for his having the concept of God is that it was implanted in him by God himself, as the mark of the maker stamped upon his work (Descartes 1996, 3.51).[11]

While genealogies of genetically justified practices can take only two possible forms, genealogies yielding justifications can take three possible forms. This is because the former turn on representations of formation processes *as being thus-and-so*, and these representations obey the principle of bivalence: they are either true or false. If a genealogy fails to yield evidence that a certain representation of formation processes is true, this will be *prima facie* evidence of its falsity. Hence, genealogies of genetically justified practices will be either vindicatory or subversive. But if we start at the other end—with the formation processes of concepts, values, and practices—we get three rather than two possible argumentative structures. This is because concepts, values, and practices are neither true nor false. They can be evaluated: there can be reasons for or against living by those practices rather than by possible alternatives. Yet this allows for the possibility that aspects of our ways of going on will simply be rationally contingent (and even where our having *some* form of a conceptual practice is not rationally contingent, the specific form it takes in our own cultural situation may be). Hence, a genealogical explanation of how we came to live by a given conceptual practice may yield reasons in favour of it, reasons against it, or neither—but as Wittgenstein (2009, §289) pointed out, the fact that we use something without justification does not mean that we use it wrongly.

At the level of these highly general considerations that are not yet specific to pragmatic genealogies, we thus find that there are two ways of connecting origin and justification that sidestep the genetic fallacy objection. Either claims about the origins of practices affect the space of reasons because these claims concern practices whose authority is itself a function of their formation, or the formation processes themselves are in some way *reason-giving*, providing reasons for or against cultivating a conceptual practice.

This brings us to an important complication we have so far ignored, namely that whether one considers a genealogy vindicatory, non-vindicatory, or subversive crucially depends on what one is prepared to *recognize as a reason* for or against a conceptual practice, and that in turn depends on one's conception of one's *ultimate aim* in assessing the ideas one lives by. Do we take ourselves to strive for the set of ideas that is absolutely and definitively best, which is to say: best from a point of view that is as free of contingent historical perspective as possible? If so, then in order for something to *count* as a reason for or against

---

[11] See Williams (2005b, 134–7).

cultivating an idea, it would have to be recognizable as a reason *to anyone*. This is to hold our ideas accountable to a timeless and universal standard—something that is by no means the preserve of believers in eternal foundations for human thought in Platonic Forms, Divine Commands, natural law, or universal reason. One can hold our ideas accountable to a timeless and universal standard while granting, at the same time, that *no ideas in fact admit* of vindication by such a standard. This is the position of Richard Rorty (1989, chs. 3–4), for example. On the last analysis, according to Rorty, all ideas must appear rationally contingent. In practice, we may be forced to retain some commitment to the ideas we happen to have; but at a reflective level, our attitude towards them should be one of ironic distance.

To assess the ideas we live by in the light of their pragmatic genealogy, however, is precisely not to hold them accountable to a timeless and universal standard; it is precisely not to ask whether *anyone* has reason to use the ideas we live by, but rather *who* does, given *which* needs and concerns. As Wittgenstein remarked: 'if anyone believes that certain concepts are absolutely the correct ones, and that having different ones would mean not realizing something that we realize—then let him imagine certain very general facts of nature to be different from what we are used to, and the formation of concepts different from the usual ones will become intelligible to him' (2009, II, §366). To see our ideas as answers to practical problems that are contingent upon who we are and what kind of world we live in is to recognize contingency not only at the level of our ideas, but also at the level of the *standards* to which they answer. This is why Williams can reject Rortyan irony as expressing a failure to go *far enough* in recognizing contingency:

> Once one goes *far enough* in recognizing contingency, the problem to which irony is supposed to provide the answer does not arise at all.... Precisely because we are not unencumbered intelligences selecting in principle among all possible outlooks, we can accept that this outlook is ours just because of the history that has made it ours; or, more precisely, has both made us, and made the outlook as something that is ours. We are no less contingently formed than the outlook is, and the formation is significantly the same. We and our outlook are not simply in the same place at the same time. If we really understand this, deeply understand it, we can be free of what is indeed another scientistic illusion, that it is our job as rational agents to search for, or at least move as best we can towards, a system of political and ethical ideas which would be the best from an absolute point of view, a point of view that was free of contingent historical perspective.   (2006d, 193–4)

Responding to this passage, John Cottingham complains that 'there is no real harmony here, just a concatenation of contingencies'; 'this is something we can perhaps learn to put up with', he concludes, 'but *confidence* seems sadly out of place' (2009, 37). *Pace* Cottingham, however, there *is* harmony between us and

our ideas when our ideas answer to *our problems*—that is the sense in which genealogy can reveal the formation undergone by us and our outlook to be *significantly* the same, and strengthen our confidence in those ideas as a result.

Of course, this kind of reflection can also weaken our confidence in our ideas by showing that they answer to problems we no longer face, or, more alarmingly, that others face in oppressing us. So Williams is not plumping for a complacent conservatism, as Srinivasan (2019, 139) might perhaps be taken to suggest.[12] Nor, as Srinivasan definitely suggests, is he disputing that the insight into the contingency of our ideas is an important one that should affect our view of them and of their relation to the ideas of others.[13] On the contrary, the crucial question becomes precisely: what are these ideas contingent *upon*? Genealogical reflection along these lines can help us locate our own ideas in relation to rival ones: are these rival ideas simply archaic, having survived into a world in which they are bereft of their point, or are they still rooted in live concerns that we merely happen not to share? And with regard to our own outlook, genealogical reflection can render us less hostage to a picture of that outlook as a tensionless and universally beneficial whole.[14] By tracing ideas to the concerns from which they derive their point, we become more disposed to recognize how many ideas only cater to certain constituencies, how one set of ideas can be deployed against another, and how the concerns they each promote can come into conflict.

On this view—which is, if not *implied* by the method of pragmatic genealogy, certainly a natural fit for it—we want the concepts and values that best make sense of the world *to us* and that best help *us* to live; but what makes sense to us and helps us to live is in turn a function of who we are, which concepts and values shape our concerns, and what kinds of circumstances we find ourselves in—all of which are largely matters of contingent historical forces. This is not a constraint to be overcome, but rather what enables our sense-making and practical reasoning in the first place. The self that subjects its concepts and values to genealogical scrutiny cannot be separated from everything that it contingently is—it is not, in the first instance, *biased* by historical processes, but *constructed* by them.[15] Once we recognize this, we can see that our task as pragmatic genealogists is not to find the ideas that are best from a point of view that is maximally pure of contingent influence, and we shall accordingly be freed of the expectation that

---

[12] Williams explicitly insists—as Srinivasan (2019, 139n19) duly acknowledges—that this does not leave us with 'an inactive or functionalist conservatism that has to take existing ethical ideas as they stand', but rather enables a 'critique of existing institutions, conceptions, prejudices, and powers' (2005h, 36–7). See Queloz and Cueni (forthcoming) for a more extensive discussion, which draws on the above, of how Williams's 'Left Wittgensteinianism' promises non-foundationalism without conservatism.

[13] See, for instance, Williams (2006d, 195).

[14] See Williams (2005h, 36–7; 2006d, 195), Prescott-Couch (manuscript), and Queloz and Cueni (forthcoming).

[15] See Williams (1993, 158–9).

our ideas ought to emerge from genealogical scrutiny vindicated against all possible rivals, in terms of reasons recognizable to anyone. The concern animating pragmatic genealogy is not that our ideas should be ultimately and timelessly desirable, but that they should have a point *for us*.

9.2 Understanding Pointfulness and Avoiding Continuity Failures

As the case studies of Chapters 4–8 have brought out, what renders formation processes reason-giving in *pragmatic* genealogies is the notion of a *point*, which straddles the space of causes and the space of reasons: pragmatic genealogies can yield reasons by showing that certain ways of going on are rational because pointful responses. But as section 9.1 also made clear, the needs and concerns that render ideas pointful themselves have a history, and the second hurdle for pragmatic genealogy's claim to being normatively significant is that this historicity threatens to rob genealogies of the continuity on which their normative significance depends. Before confronting this second objection head on, however, we first need to clarify the notion of 'pointfulness' at work in pragmatic genealogies. What exactly is it for an idea to have a point?

9.2.1 The Need-Satisfaction Account of Pointfulness

Ideas—more precisely, tokens of a conceptual practice—have countless effects. So what makes a particular type of effect the *point* of an idea? Evidently, pragmatic genealogists' talk of points and pointfulness is really a humanistic-sounding way of ascribing *functions* to conceptual practices, and in the literature on the notion of function, there are broadly speaking two kinds of theories on offer: *causal role* theories and *selected effect* theories.[16] The first kind of theory focuses on the way in which talk of functions helps us understand the causal role of something within a complex system: by asking which effects of some particular trait or item contribute to the realization of some system-level capacity—which effects of the boiler contribute to the heating system's capacity to heat the house, for example—we can specify the relevant effects of something in relation to a system-level capacity. The second kind of theory, by contrast, focuses on the way in which talk of functions helps us understand how the past effects of something contributed to its present ubiquity: by asking which effects have a history of being selected for— which among the various past effects of hearts (heat production, noise emission,

---

[16] The first kind of theory is exemplified by Cummins (1975); the second by Wright (1973), Millikan (1989), and Neander (2017).

or blood pumping) are causally responsible for the retention and consequent prevalence of hearts, for example—we can specify the relevant effects by reference to selection histories. While the first kind of theory has been criticized for being too inclusive in what it is prepared to describe as a function, the second has been criticized for being too restrictive in limiting functions to effects with a selection history.[17]

What selected effect theories have going for them is that they make ascriptions of functions more objective. The importance of this is brought out in psychiatry, for instance, when attempts are made to define mental illness in terms of functions and dysfunctions in the brain.[18] Who gets to decide what counts as a functional brain? Are ascriptions of functions just thinly veiled value judgements? By equating functions with selected effects, we can do away with much of this dependence on value judgements. There is an objective fact of the matter as to what the selection history of a trait looks like, so that one can discover functions largely independently of one's value judgements.

Insofar as one's concern in telling a pragmatic genealogy is the purely explanatory one of accounting for the present ubiquity of an idea, the relevant understanding of pointfulness arguably aligns with selected effect theory. But insofar as the driving concern of our pragmatic genealogists is to find out whether an idea is *worth having going forward*, understanding the points of conceptual practices as selected effects will not do. Conceptual practices may now serve an important need *even if* that need played no role whatsoever in the retention of that concept in the past. A conceptual practice could be freshly instituted by a mad king on a whim—so that on a selected effect theory, it would lack a point—while still promising to stand in important instrumental relations to our needs going forward.

For the purpose of understanding the pragmatic genealogies at issue in this book, therefore, the notion of pointfulness is better understood not in terms of selected effects, but in terms of *needed* effects: the points of conceptual practices are the effects that contribute to the satisfaction of the needs of concept-users. More precisely, to say that the point of concept A is to serve some need C is to say:

(1) Tokens of the practice of living by concept A tend to cause effect B.
(2) B tends to cause the satisfaction of need C.

(1) and (2) show how ascriptions of pointfulness or functionality can be translated into causal claims relating the typical effects of conceptual practices to the satisfaction of human needs. We can call this the *need-satisfaction account* of

---

[17] For an overview of the different theories of functions in biology and the criticism they have attracted, see Garson (2016).
[18] See Garson (2019, ch. 11).

pointfulness or functions. On this account, the notion of pointfulness at work in pragmatic genealogies is a fairly modest causal notion which is not committed to there being any particular selection mechanism or even any kind of differential survival—(1) and (2) would be applicable even to a world containing little else besides *A*, *B*, and *C*.[19] The notion of pointfulness they articulate is simply the idea that the practice of living by a given concept, value, or virtue has certain effects, and that some of these effects make a *salient useful difference* to the lives of participants in the practice, where 'usefulness' is specified in terms of conduciveness to need-satisfaction, and 'saliency' is specified according to the purposes animating the telling of the genealogical story in the first place: by describing something in terms of its point, one highlights a select few among the unsurveyably many effects of a practice, and the merits of the selection depend on the purposes pursued in so describing them.[20]

How perspectival or mind-dependent does this make the functions or points of ideas? Certainly more so than on selected effect theories, where functions are objectively determined by causal histories, and the selection patterns discernible in those histories are, in that sense, formed independently of human interests. But while severing the connection to human interests secures greater objectivity, it also renders it less evident that *this* should be the notion of function best suited to assessing an idea's present relation to our interests. In the context of evolutionary biology, there is certainly a point to thinking of functions as selected effects, but that point derives from the professional interest of evolutionary biologists in selection histories. When our interests are not primarily historical, however, it is less clear that this is the concept of function we need (a fact obscured by the frequent overlaps between the effects that different interests give one reason to highlight—the evolutionary biologist and the heart surgeon may approach the heart with very different concerns and nonetheless converge in the effects they find worth highlighting). We must apply the spirit of pragmatic genealogy to its operative concepts.

Unlike selected effect theories, causal role theories retain a tight connection to human interests: by specifying the function of something in terms of how it contributes to the realization of some system-level capacity, we identify its function in relation to *what we see as a system* exercising *a capacity we are interested in*. Here the human disposition to regard something as a system worth understanding is primary and grounds function attributions.[21] However, this also means that things will have as many functions as there are systems we are interested in: relative to the system of honey production, the function of bees is to produce

---

[19] A point highlighted by Kincaid (2020, 23) to immunize functionality ascriptions against the notorious 'missing mechanism objection'. For a critical discussion of this objection, see Van Riel (2020).

[20] See Queloz (2019) and Barnes (1995, 43).   [21] See Haslanger (forthcoming, §6).

honey; but relative to the system of absinthe production, the function of bees is to pollinate wormwood. Hence the charge that causal role theories are overly inclusive.

On the need-satisfaction account, by contrast, we do not have to start from the idea that there is some system whose capacities we want to explain in terms of the capacities of its constituents; nor are we required to specify in what sense a community of concept-users form a 'system' exercising some 'system-level capacity' before we can attribute functions to our concepts; what guides the attribution of functions is rather the conviction that we have certain needs relative to which certain causal effects of conceptual practices are usefully singled out because they contribute to the satisfaction of those needs. Function attributions become less arbitrary when they are anchored in needs rather than in what we are prepared to regard as a system. We identify functions by working from the needs up rather than from the system down.

Of course, the ascription of needs still involves interpretation and value judgements (as, in a different way, does the individuation of systems and capacities). But need ascriptions are still more objective than many other value judgements: whether we have a need comes as a discovery, and it is not subject to our will the way that having a desire, a preference, a purpose, or a goal is.[22] We can have needs without knowing that we have them, and part of what pragmatic genealogies do is to help us recognize the needs we did not know we had by deriving them from needs we knew we had.

If we apply the spirit of pragmatic genealogy to the notion of pointfulness it operates with, it becomes clear that how inclusive its understanding of functionality can properly get once again depends on what we use it for. In theories of biological functions, a notion of function will arguably be too liberal if it includes so much that it ends up being out of touch with the way biologists think and speak.[23] But in the present book, the task is to make sense of the notion of functionality or pointfulness as it figures in the pragmatic genealogies we considered. And the need-satisfaction account does achieve that: each of the pragmatic genealogies we considered primarily turns on revealing the way in which some conceptual practice satisfies some important need, be it the need to avoid conflicts over external goods (Hume), the need to avoid deception within the community (Nietzsche), the need to flag good informants (Craig), the need to gain

---

[22] This helps distinguish the need-satisfaction account from accounts on which functions are, more broadly, effects that promote the realization of people's goals: see Wimsatt (1972), Boorse (1976), and Nagel (1977); more recently, the idea that functions are contributions to goal-realization has been defended by McShea (2012), Trestman (2012), and Piccinini (2015, ch. 6). A related account that has been influential in discussions of social functions is Searle's (2010, 58–60), on which functions are imposed on objects by agential purposes and values.

[23] See Garson (2016, chs. 1–3).

and share information effectively (Williams), or the need to neutralize prejudice (Fricker).

In the first instance, of course, these are ascriptions of pointfulness to proto-types of conceptual practices in relation to the needs of agents in a state-of-nature model. The genealogies primarily identify instrumental relations *within a model*, and it is important to distinguish this primary use from secondary uses that build on it, such as drawing attention to comparable instrumental relations in our actual practices, or making out comparable instrumental relations in the history of our actual practices. Uncluttered by the messiness of reality, the model sharpens our eye for certain patterns of pointfulness (much as a priming look at a sample morel can help one spot the notoriously well-camouflaged morels of varying shapes and colours in the tangles of the forest floor). The model also provides *prima facie* evidence for ascriptions of pointfulness in much the same way that design analyses in evolutionary biology provide evidence for ascriptions of biological function: a model is used to show that a given trait would solve a problem, and this is advanced as evidence for thinking that what we find in reality solves a comparable problem.[24]

Insofar as such instrumental relations can plausibly be identified in reality—either now or in the past—they can then act as a basis for *evaluations* of the extent to which we have reason to continue to engage in the practice, and they can act as a basis for more or less ambitious *explanations*. For example, instrumental relations can act as a basis for explanations of *resilience* (Pettit 1996, 299–300), i.e. of why a practice is in some respects unlikely to disappear, because its loss would make itself felt in ways that would drive the practice back into the mainstream. This is to undertake, in addition to the commitments to claims (1) and (2), a further commitment to the following claim:

(3) *A* is resilient because it tends to cause the satisfaction of need *C*.

This implies a commitment to *counterfactual* claims to the effect that certain forces *would* be actualized *if* we were to move away from *A*. Even more ambi-tiously, instrumental relations can act as a basis for explanations of *persistence*, i.e. of why a practice actually endured over time:

(4) *A* persists because it tends to cause the satisfaction of need *C*.

This implies a commitment to *factual* claims to the effect that such forces *were* actualized and are part and parcel of the causal-historical story explaining why we now find *A*. But as Harold Kincaid (2020, 21–2) makes clear, even such

---

[24] See Kincaid (1996, 118–19) for a discussion of design analyses in biology. I elaborate on the differences between pragmatic genealogies and design analyses in Queloz (2020).

explanations of persistence can be cashed out in terms of unmysterious causal claims of the form: the existence of practice A at time t can be explained by the fact that the typical effects of A at time $t - 1$ cause the existence of A at time t. Neither explanations of resilience nor explanations of persistence involve the claim that a practice *came into existence in order to* serve human needs (a claim that would invite the old objection that as yet unrealized effects cannot cause something to exist). They only explain why, having come into existence, a practice is unlikely to disappear (resilience) or why, having come into existence, it was retained (persistence). Only the use of pragmatic genealogical models in explanations of persistence carries any claims about the actual course of history. Using such models merely to ascribe pointfulness or resilience to our present practices does not yet commit one to history being a certain way.[25]

What commitments and burdens of proof are undertaken by pragmatic genealogies thus depends on the use to which they are put by the genealogists and their audience. This methodological nuance is registered most clearly by Craig: 'The depth of factual obligation incurred by a state-of-nature theory depends on its aims', he writes; it 'will be greatest when its intentions are explanatory, to account for the existence of the target phenomenon' (2007, 193). By contrast, the depth of factual obligation will be smallest when the story is offered merely as a heuristic device that helps us determine to what extent certain instrumental relations now obtain between needs and conceptual practices.

How does Craig's own genealogy situate itself on that spectrum? Considered in isolation from the declared aims of its author, the genealogy minimally involves an ascription of pointfulness to the concept of proto-knowledge in the model: the practice of living by the concept of proto-knowledge causes the flagging of good informants; the flagging of good informants helps satisfy the need to pool information; and therefore, the practice of living by the concept of proto-knowledge-helps satisfy the need to pool information. But these observations might then be used as a basis for the identification of similar instrumental relations in our *actual* conceptual practice. And having identified these relations in our actual practice, they might further be used to explain why, were the practice of living by the concept of knowledge to come under pressure, there would be some pressure to drive it back into use. Or they might be used to explain why the concept of knowledge persisted up to the present day in so many cultures. Craig himself declares that he 'was trying to explain how certain real results have arisen, and only real pressures can produce real results' (2007, 190). Accordingly, he notes:

> I do and must suppose that there were societies whose members, collectively and individually, had the needs I ascribe to them and were able, whether as the

---

[25] See Queloz (2020) for further discussion.

outcome of some conscious process or of other equally real tendencies, to find their way to the solution I describe.... My line was, and had to be, that the needs were real and the persons concerned would have come, in one way or another, to satisfy them.... I had to maintain that the circumstances that favour the formation of the concept of knowledge still exist, or did until very recently, since otherwise I would have had no convincing answer to the obvious question why it should have remained in use.   (Craig 2007, 191)

This means that Craig's ambitions in advancing his genealogy led him to incur rather deeper factual obligations than he would have incurred had he simply used the genealogy to reveal the relation of the concept we now have to some of our present needs. And one might complain that his book, admirably concise though it is, marshals rather little empirical evidence with which to honour these obligations. But one upshot of the present discussion is that the merits of his genealogical model are distinct from the merits of the use to which he puts it. We can find fault with a tool's application while thinking no less of the tool.

9.2.2 Avoiding Continuity Failures

Having clarified what it means to identify the pointfulness of a conceptual practice in some situation of emergence, we are now in a position to confront the second objection to the claim that pragmatic genealogies can affect the space of reasons: showing that something has a point in some situation of emergence does not suffice to show that it now has one, since the conditions from which a practice originally derives its point may not obtain in the situation we are now in—the genealogy's normative ambitions might founder on what Nicholas Smyth calls *continuity failure* (2017, 1137).[26]

This second objection thus points to the fact that if they are to possess normative import, pragmatic genealogies presuppose continuity in the conditions relative to which a practice has a point. We can formulate this constraint as follows:

Continuity:

Necessarily, for any P, G, and RC_i: if $\{RC_1, RC_2, \ldots RC_n\}$ is the set of root conditions relative to which practice P is originally pointful under some description, then the inference from the original point of practice P to its actual point in group G is justified only if $\{RC_1, RC_2, \ldots RC_n\}$ also obtains in G.

---

[26] Smyth's (2017) critique concerns inferences from the original to the current function of morality, and he focuses on genealogists such as Kitcher (2011) and Joyce (2006), though there is a suggestion that the same difficulties extend to Hume and Williams—see Smyth (2017, 1130n4, 1131).

When this constraint is not met, genealogically derived ascriptions of pointfulness to actual practices suffer from continuity failure. If we are to draw any evaluative conclusions about our practice from its practical origins depicted in the genealogical model, therefore, we must be given some reason to think that there is in fact continuity between the conditions that give point to the practice in the model and those we actually face.

There are two strategies with which pragmatic genealogies can try to avoid continuity failures. The first is to operate at high levels of description that abstract away from particulars and bring into view features extending over a wide range of situations. Call this the *high level of description* strategy. The second is to show that the need for the target practice has a firm basis in and derives from basic needs humans can be assumed to have anyway. Call this the *anchoring in basic needs* strategy. While analytically distinct, the two strategies are combined in the pragmatic genealogies we considered: they focus on the general and anchor it in the basic.

The danger for this way of proceeding is that the explanations will end up being too abstract and general to be informative. It is therefore no coincidence that the pragmatic genealogies we have encountered take a *piecemeal* approach: instead of trying to identify the point of entire domains of human thought and action, they proceed *one conceptual practice at a time*, singling out a particular thread within the tangle of our conceptual practices and following it to its moorings in the needs of concept-users. Though this may not be a necessary condition on the method's effectiveness, it contributes to it in two ways. First, since informativeness decreases with increasing abstraction, but securing continuity requires working at a fairly abstract level, keeping the object of investigation narrow and concrete by philosophy's standards—showing that any society will need a *particular* conceptual practice in order to solve a *specific*, well-delineated coordination problem, for instance—allows one to maximize informativeness while retaining continuity. Second, working piecemeal keeps one more sensitive to the extent to which our practices are an assemblage of individually pointful elements, each tailored to its specific point, that do not all fit together into a harmonious, functional whole. If one does not work piecemeal and inquires into the point of *morality* rather than of a particular moral idea, one is more likely to miss the tensions and conflicts between ideas that cannot be pursued all the way together.

It is true that even if one works piecemeal, substantial commitments will be undertaken about what kinds of creatures humans are and what kinds of environments they live in. Like all explanations, genealogical explanations have to start somewhere and take certain things for granted. But the pragmatic genealogists do not simply settle on a practice and then paint an innate need for just that practice into our picture of human nature. They execute their genealogical projects in a way that allows them to take as little for granted as possible. They try wherever possible to take for granted only *structural needs* such as the need to gather and

share information about the immediate environment or the need to avoid conflict: as we saw in Chapter 4, structural needs are second-order needs that grow out of the relations between individuals' first-order needs (such as their need for various types of foods, goods, and tools). Structural needs are to a large extent counterfactually robust, because they are overwhelmingly insensitive to the content of first-order needs: almost irrespective of what their first-order needs are under given circumstances, humans will have a strong interest in gathering and sharing information about their immediate environment and in avoiding conflict. Indeed, there is a limit to how different a form of life can be while remaining intelligible as a variation on ours. Differences must ultimately be made sense of in terms of similarities—variations on human life are only recognizable as such against the backdrop of a shared set of features that make them variations on human life. And if there are any needs that we can be confident human beings have anyway, structural needs that grow already out of the least contested of human needs, such as the need to locate and access sources of water, are good candidates.

These structural needs can then form a basis from which to derive further needs, by showing how one practical exigency entails the next, until one reaches the need to which the target practice forms a direct response. The need for the target practice is thus not simply stipulated, but shown to be entailed by less controversial needs. This way of proceeding sets pragmatic genealogies apart from much-maligned just-so stories about human nature such as that of Randy and Nancy Thornhill (1983, 1992), which raised eyebrows by presenting 'men's tendency to rape' as an innate part of human nature.[27] Pragmatic genealogies are more modest in their assumptions when they demonstrably but fallibly *derive* the needs they are interested in from structural needs that raise no eyebrows when presented as inscribed in human nature.

The question raised by *Continuity* is whether we actually have the needs at issue in the genealogical model, and there is a point to the *anchoring in basic needs* strategy when target needs we are *less* confident we have can be derived from root needs we are *more* confident we have. It may not at all be obvious that we need the virtues of truth, the concept of knowledge, or the virtues of justice in the Humean and Nietzschean senses. What pragmatic genealogies do is reveal how some practice helps us to live by taking something we are less confident we need (e.g. the concept of knowledge, or the virtues of truth) and deriving it, as a practical corollary, from something we are more confident we need (e.g. information about our immediate environment). Eschewing attempts to derive the concepts we should live by from absolute rational foundations in universal reason, they instead

---

[27] See Hufendiek (forthcoming) for a nuanced discussion of this and other controversial inscriptions of traits into human nature and of how these have been exploited as bases for the critique of naturalism in the nature–nurture debate.

try to *foster* allegiance to certain conceptual practices by showing that they promote material that *already* commands allegiance. The uncontroversial needs that figure at the root of the pragmatic genealogies are paradigmatic examples of such material. Few will be disposed to deny that we have these needs; what they *might* be disposed to deny is that these needs bring with them certain problems that certain conceptual practices in turn equip us to solve; and this is where the genealogical derivations come in, as narratives designed to bring out just how these conceptual practices in fact serve ends that the narrative's addressees are already fully committed to pursuing (which is not to say that the most basic needs are always those that command the most allegiance or that we are most confident in—the diehard liberal may be willing to sacrifice a great deal before compromising on the need for political liberty).

The crucial point, then, is that pragmatic genealogies do not *assume* continuity in the practical demands we face, because precisely what they are is attempts to identify *bases* of continuity in those demands. They are not arguments depending on continuity, but arguments *for* it. The argumentative structure of a pragmatic genealogy can be reconstructed as follows:

(P1)   In a prototypical group G, a set of root needs RN_1–RN_n under root conditions RC_1–RC_n generates a practical problem.

(P2)   This generates a practical pressure on G to solve the problem: the target need TN.

(P3)   Prototypical conceptual practice CP would meet the target need TN by serving point P.

(P4) CP could develop quite naturally, i.e. out of the capacities we are prepared to grant G anyway, via the set of steps S_1–S_n.

(C1)   Therefore, circumstances permitting, CP would be highly likely to develop in G.

(C2)   Therefore, it is rational for G to engage in CP in order for P to be served in G (in the sense that people with these needs under these circumstances would welcome and, if they could do so, aim for engagement in CP with a view to securing P).

(P5)   In the actual group $G^\star$, there are close analogues to RN_1–RN_n and RC_1–RC_n, namely $RN^\star_1$–$RN^\star_n$ and $RC^\star_1$–$RC^\star_n$.

(C3)   Therefore, it is also rational for $G^\star$ to engage in $CP^\star$, the closest analogue to CP in $G^\star$, in order for P to be served.

(C4)   Therefore, the best explanation for why we go in for $CP^\star$ is that it serves P.

(C5)   Therefore, there is a *prima facie* reason for $G^\star$ to continue to engage in $CP^\star$, and $CP^\star$ is to that extent vindicated.

This reconstruction lays out how pragmatic genealogy can affect the space of reason by showing us that *given* that we share certain needs, we have reasons to engage in certain conceptual practices.[28] What the reconstruction also brings out is that the soft underbelly of such genealogies is (P5), which assumes that the root needs and root circumstances *in fact obtain* in our present situation. It is this premise that the two strategies we considered aim to strengthen: the variables RN^*_1–RN^*_n and RC^*_1–RC^*_n are assigned to facts about human beings and their environment that stand a good chance of obtaining anyway, independently of the particulars of a given situation, because they are basic structural facts about the human situation picked out under highly general and abstract descriptions.

A pragmatic genealogy thus aims to affect the space of reasons through an inference from a generic predicament to our local manifestation of it. Such an inference might still be wrong, of course. Yet on the interpretation offered here, the problem will then not be that it has subtly trespassed against the canons of reasoning, but simply that it is unsound.

9.3 The State of Nature as a Model of Local Problems

Emphasizing the respects in which pragmatic genealogies can avoid continuity failures by building their case on near-universal needs captures a central concern of the genealogical projects we considered: to bring out the respects in which some of our conceptual practices respond to timeless human problems. But it also invites the objection that this severely restricts the explanatory scope of the pragmatic genealogical method. It suggests that the method is appropriate only when dealing with anthropologically necessary conceptual practices—and surely the greater part of human thought is not necessary in that way.

While many of the pragmatic genealogists we considered are indeed keen to show that certain conceptual practices are, at core, enduring because held in place by near-universal human needs, we must be careful not to mistake incidental for essential features of the method. In particular, we should not take pragmatic genealogy to be limited in principle to investigating what P. F. Strawson called the 'central core of human thinking which has no history—or none recorded in histories of thought' (1959, 10). The response to the threat of continuity failure explored in section 9.2 was, in effect, that the genealogical model applies to *us* because it applies to *any* human beings anywhere. But this is only *one* way of

---

[28] A complication arises here—at least on an internalist account of reasons—if we acknowledge that someone might lack any interest in pursuing their needs, where this lack of interest is not the product of a false belief, and there is no sound deliberative route, from the motives that they do have, by which they might be brought to care; but as Williams notes, 'we have to bear in mind how strong these assumptions are, and how seldom we are likely to think that we know them to be true' (1981b, 105).

securing continuity between the model and reality. The other way is for the model to apply to us because the conditions it models are specifically *ours*.

Balancing out section 9.2's one-sided emphasis on the near-universal, this section therefore emphasizes that pragmatic genealogy is neither constitutively committed to there being an *enduring core* at the centre of the practices it investigates nor restricted to considering only *universal* needs. A significant upshot of the interpretation defended here is rather that pragmatic genealogy can be tailored to our specific situation by modelling even highly local problems arising from local needs. Contrary to its history-transcending connotations, the state of nature can help us make sense of our own particular location in history and of its relation to other socio-historical situations.

It is true that Craig in particular tends to present his method as revealing 'the core of the concept as it is to be found now' (2007, 191), a core he presents as 'an outcome of certain very general facts about the human situation' (1990, 10). He thereby commits himself to the thesis that the prototype of the concept of knowledge described in his genealogy makes up the core of our actual conceptual practice. But it would unnecessarily weaken the method to view that commitment as essential to it. As Kusch and McKenna rightly insist, 'we should not think of *protoknowledge* as the core or essence of *knowledge* just because we have a predictively successful model that represents *knowledge* as developing out of *protoknowledge*' (2020, 1061). Part of the power of pragmatic genealogy is that it can make sense of our conceptual practices as elaborations of simpler prototypes even when these prototypes are not themselves realized in our actual practices. It can help us make sense of practices as elaborations of prototypes that are no longer—and perhaps never were—extant.

Pragmatic genealogy is thus not in principle committed to there being an unchanging, timeless core at the centre of the practice it investigates, even if some genealogists encourage that preconception. It will therefore not fall quite so easily into what the historian Peter Gay dubbed 'the trap of spurious persistence' (1971, 192)—the mistake of treating ideas as more unchanging and insensitive to historical context than they really are.[29]

Harder to dislodge is the assumption that the state-of-nature model necessarily depicts universal human needs. Again, this is an idea that the pragmatic genealogists themselves encourage. Craig traces the concept of knowledge to some 'very

---

[29] A danger that the methodological debates initiated by Quentin Skinner, J. G. A. Pocock, John Dunn, and other figures associated with the Cambridge School have been effective in alerting us to. Skinner's diatribe against the idea that the thinkers of the past could be understood as explicating a set of 'fundamental concepts' and 'universal ideas' in answer to a set of 'perennial problems' (Skinner 1969, 5) has left its mark, engendering an acute and carefully cultivated 'sensitivity to anachronism' (Oakley 1999, 9). As a result, historians have become so uneasy about tracing 'the morphology of a given concept over time' (Skinner 1969, 5) as to put the approach explored here effectively off limits. I hope to show that on the interpretation I defend, pragmatic genealogy is more accommodating of historical change than it at first appears.

general facts about the human situation', so general, indeed, 'that one cannot imagine their changing whilst anything we can still recognize as social life persists' (1990, 10); Williams describes the state of nature as a 'representation of universal requirements' (2014g, 410), while Fricker describes it as a 'maximally ahistorical setting' serving to characterize our most basic needs and what they entail (2007, 108–9). If one takes this to be a necessary feature of the method, it must seem odd that historically, the state-of-nature method should have been most prominent in political philosophy. It is, after all, one of the more history-sensitive branches of philosophy, which concerns practices that are fairly specific to particular cultural situations in comparison to the highly generic conceptual practices that are the concern of the philosophy of language and mind. This is true even of the state, which in the wake of Hobbes is often seen as the paradigmatic object of state-of-nature theorizing in political philosophy:[30] for most of prehistory, human societies were stateless societies.[31]

Against this conception of the state of nature, it is worth emphasizing that one of the more interesting consequences of interpreting the state of nature *as a model* is that it is cut loose from the requirement of having to depict near-universal needs. Even if the genealogical method invites us to start from the most generic needs that still prove illuminating, these may still be comparatively local, and there is nothing in the method to rule out that the most generic need that a practice bears an instrumental relation to will be a recent arrival on the historical scene. Not only is pragmatic genealogy not restricted to near-universal needs—it is not even bound to *start* from near-universal needs. It is true that if one's aim in telling a genealogy is to advocate whole-hearted engagement in a target practice by presenting it as indispensable for us, then grounding a need for the practice in needs so general that one cannot imagine their changing while human life persists is an effective strategy. But more parochial material can command strong allegiance as well, and it can be just as effective to ground a need for the target practice in needs that 'we', in a more local sense, consider non-negotiable. The pragmatic genealogist can use the state of nature to model local problems that are specifically those of certain people in certain times and places. All it takes is for the state-of-nature model to be given a *localizing* rather than a universalizing interpretation—an interpretation on which the genealogy is understood to be a pragmatic reconstruction *of our particular situation* rather than of some generic human condition.

This can even be done with models that were originally intended to be given a universalizing interpretation. A particularly prominent example—which bears at

---

[30] In fact, it is arguably the concern with natural rights that constitutes the animating concern for the use of the state-of-nature device in political philosophy from Grotius through Hobbes, Locke, and Selden to Pufendorf. See Tuck (1979, 1993) and Lane (1999).
[31] See Service (1975) and Johnson and Earle (1987).

least a strong family resemblance to state-of-nature models—is John Rawls's 'Original Position'.[32] In *A Theory of Justice*, Rawls interprets the Original Position as the depiction of a timeless problem; to regard the human situation from the perspective of the Original Position 'is to see it *sub specie aeternitatis*: it is to regard the human situation not only from all social but also from all temporal points of view' (1971, 587). But in his later *Political Liberalism*, he interprets the same model as modelling a local problem, one specific to modern societies with their unprecedented moral diversity and attendant difficulties in reaching substantive agreement: the problem of providing a shared framework justifiable to people with remarkably different ethical outlooks and a correspondingly thin basis for negotiating and justifying the structures of society (1993, xviii). Outwardly, the model remains the same. But the interpretation of *what it does* and *for whom* it is supposed to have a point has changed.[33]

Another example of a pragmatic reconstruction of a local problem's comparatively generic form that nonetheless still calls for a localizing interpretation is Christoph Möllers's examination of the point of the separation of powers in *The Three Branches* (2013). The book can be read as constructing a jurisdictionally neutral model of the separation of powers by asking *of what generic prototype* the various historically and culturally inflected concretizations of the separation of powers in different national jurisdictions are elaborations. Because what we see in reality are very different concretizations of the separation of powers in different jurisdictions, working back to a neutral prototype allows Möllers to reverse-engineer the most general point of having three branches of government in the first place—it allows him to determine what the separation of powers *does for us* before it becomes entangled in local traditions. On Möllers's account, the point of separating power into three branches—the legislative, the executive, and the judicial—is to negotiate the otherwise irredeemable tension between individual and collective self-determination. The legislative and executive branch, which together serve the need for collective self-determination, are balanced against the judicial branch, which serves the need for individual self-determination. The separation of powers thus allows the tension between individual and collective self-determination to be 'perpetuated as a political controversy within its framework' (Möllers 2013, 8). But though this working back to the most general point of the separation of powers in its jurisdictionally neutral form can look like an attempt to shake off any kind of historical conditioning, it is clear that even this most general point remains relative to needs that, though widely shared across modern societies, remain indexed to a particular stretch of history. The idea of collective self-determination, in particular, is a distinctively modern idea.[34]

---

[32] See Rawls (1985) and Williams (2014d, 326–7). Rawls remarked in 1959: 'the conception of justice I set out is perhaps closer to Hume's view than any other' (Forrester 2019, 12).

[33] See Queloz and Cueni (forthcoming, §3).   [34] See Habermas (1996, ch. 3, §1).

Nevertheless, tracing the separation of powers to such local needs allows Möllers to vindicate it against the widespread suspicion that it might be a relic of an outdated legal order that has outlived its usefulness (2013, 8, 227). It also gives him a critical grip on realizations of the separation of powers, allowing him to assess the extent to which realizations of the separation of powers in particular jurisdictions serve or fail to serve that point (2013, ch. 3); and it gives him some guidance in determining whether there might be a point in trying to replicate such a separation at the international level (2013, ch. 4). Working back to the jurisdictionally neutral prototype filters the practical contribution that the separation of powers makes at the generic level from the ways in which particular concretizations of it add to—or detract from—that contribution.

A third example of the pragmatic reconstruction of a local problem's comparatively generic form is Damian Cueni's (2020) account of what he calls the *Liberal Democratic Dilemma*. Taking his cue from Craig, Williams, and Fricker, Cueni works back to the most generic form of that dilemma (2020, ch. 1) before considering its manifestation and elaboration in more specific socio-historical contexts (2020, ch. 5). But as calling it a *liberal democratic* dilemma acknowledges, the problem remains rooted in local needs characteristic of liberal democratic societies. The dilemma is that we who live in liberal democratic states face two conflicting demands: on the one hand, we increasingly need *strengthened international institutions* to tackle the pressing problems of our time (climate change, financial systemic risk, health security, the global migration crisis); on the other hand, we still need the power exerted by these international institutions to be *legitimated according to liberal democratic standards of legitimacy* if that power is to be reconciled with our modern concern for individual and collective self-determination. Traditionally, in the domestic context, reconciling our need for rule with our need for legitimacy has been made possible by conceptual and institutional resources allowing us to mark out certain exercises of power *as* legitimate, and in liberal democratic states, these resources have taken the form of democratic participation, accountability, transparency, and judicial review. But since these resources are tailored to the legitimation of institutional power *within* the nation state, they are notoriously ill-suited to legitimating the very different forms that institutional power takes *beyond* the nation state. As a result, we are caught up between the conflicting demands of two real needs: the need to strengthen international institutions on the one hand, and the need to legitimate their power according to our domestic expectations of legitimacy on the other. The effective tackling of pressing global problems seems to come at the expense of collective and individual self-determination; the preservation of collective and individual self-determination at the cost of a failure to address pressing global problems. Historically, this predicament is extremely local. But this makes it no less pressing and ineluctable for us.

This suggests a more general methodological point. The mere fact that a pragmatic genealogy starts out from historically contingent problems renders the solutions no less necessary. All that the insight into the contingency of the problem entails is the theoretical possibility of ridding oneself of the problem by eliminating the features of our situation that engender it. But as the example of global challenges and international institutions illustrates, this will often be no more than a theoretical possibility. Certain conceptual practices may turn out to be indispensable for us even if they are the products of historically contingent circumstances. In these cases, pragmatic genealogy will have uncovered what might be described as indispensability without inevitability.

The upshot of this model-based, progressively localizing interpretation of pragmatic genealogy is that we are not faced with an exclusive choice between historical genealogy and model-based genealogy; between a method concerned with the socio-historically local and a method concerned with the maximally ahistorical. The state of nature and its functional equivalents can represent demands that are to various degrees socio-historically situated. To put it cursorily: the state of nature can model local problems, and situate them in relation to generic ones. Pragmatic genealogy is thus not committed to viewing our local situation only under its most universal aspect as the latest iteration of the condition of generic humanoids.

9.4 Contested Needs and the Conception of the Agent

Having established that neither the *explananda* nor the *explanantia* of pragmatic genealogies have to be anthropologically necessary—such genealogies are neither restricted to explaining the universal core of human thought nor to doing so in terms of universal needs—we are left with the last of the four objections we set out from: needs do not provide a solid basis for genealogical explanations, because need ascriptions are contestable. What seems like a need to some seems to others like a mere caprice.

So far from ignoring the fact that need ascriptions are contestable, however, pragmatic genealogies give us the means to confront such contestations. Pragmatic genealogies are tools by which to ascertain whether we rightly treat something as need. Not only can they show that we need things we may not have suspected we needed; they can also help us ground controversial need ascriptions in less controversial need ascriptions—at the limit, revealing a need to be ineliminable because rooted in needs we have anyway. They help us relate things we are *not sure* we need and ought to value 'to other things that we *know* that we need and value' (Williams 2002, 90, emphasis mine). As Craig writes, genealogies are 'at their strongest when the human needs from which they start are the most practical, hence the most undeniable ones' (1990, 89). But of course, even the

most undeniable needs have been denied. The idea of a need is correlative with the idea of a serious harm that one will incur if the need is not satisfied, and that idea of a serious harm is in turn correlative with culturally conditioned conceptions of human life and flourishing. These ideas are contestable, but as David Wiggins and Sira Dermen have argued, 'that is the condition of all important ideas' (1987, 63), and they are none the worse for that.[35]

If pragmatic genealogies tend to ground contested need ascriptions in further need ascriptions, it is because needs are in important respects more objective than desires, preferences, aims, or purposes. Most basically, what one needs does not depend on the workings of the mind in the way that it does in those other cases. Needs are objective in that they are not subject to the will: we cannot, in the relevant sense, simply *decide* to need something. To need X because it is F, moreover, X must really *be* F, while one can desire X because it is F even when X is not F. Furthermore, what our needs are is not necessarily transparent to us in the way that our individual wants and purposes are transparent to us—in modern parlance, needs are not *luminous* (Williamson 2000, 13): we may have them without knowing that we have them, and we can think we have them without really having them. Moreover, needs could be said to be objectively demanding, making demands on us in a way that mere desires and purposes do not—not just practical demands on those whose needs they are, but ethical demands on others. As Williams puts it in an early essay, it is 'a matter of logic that particular sorts of needs constitute a reason for receiving particular sorts of good' (1973c, 241–2), whereas it is much less clear that the same could be said of interests, merits, preferences, or purposes. This combination of features contributes to turning the question of what our needs *really* are into a moot question that carries substantial ethical and political implications.[36] We have, for example, rightly become wary of ascriptions of *real* as opposed to *perceived* needs, on the grounds that such ascriptions might be used to coercively override people's perceived needs. Down this road, the fear is, lies totalitarianism. But note that the problem lies not in the notion of real needs itself, i.e. in the idea that there are such things. The problem lies in a further idea, namely the principle that licenses certain inferences *from* the ascription of real needs *to* certain practical and political conclusions—in particular, that real needs justify coercion.[37]

The need ascriptions of pragmatic genealogies anyway do not hinge on the contrast between real and perceived needs, but rather on that between needs we *already* perceive and needs we can be *brought to* perceive. These genealogies concern needs we may be unaware of or insufficiently sensitive to, but which we

---

[35] The notion of a need is explored in great detail in Wiggins (1998, 2002a, 2005).

[36] See Brock (1998), Reader (2005, 2007), and Brock and Miller (2019) for further discussion of the normative significance of needs for ethical and political theory.

[37] See Williams (2011, ch. 3).

can come to perceive through a perspicuous representation of how these needs are entailed by needs we already perceive ourselves as having. They offer derivations of needs we did not know we had from needs we knew we had. The process whereby we come to see needs we did not see before is not problematically self-validating in the way that, say, brainwashing would be. No controversial idea of needs is at stake in pragmatic genealogy if all it is committed to is that there are needs that, through genealogical reflection, we can be brought to recognize in light of needs we acknowledge already. And even when what acts as the point of departure in a genealogical narrative is not beyond doubt, the method can be applied to this point of departure in turn. We can first reflect on one part of our conceptual practices while relying on the rest, and then take *that* part for granted in sounding out some other part. In good Neurathian fashion, we mend the ship while out on the open sea.

What is true, however, is that although needs are objective in many respects, they are so only relative to the *conception of the agent* one implicitly draws on in telling a genealogy. In ascribing needs to the agents in one's initial description of the state of nature, and less obviously also in moving from this to later stages of the story, one draws on a particular conception of the kind of creature we are dealing with. The conception of the agent with which pragmatic genealogies operate is itself an important parameter in those genealogies, and one that can be made to take different values in the course of the genealogy. A pragmatic genealogy thus only ever shows to what extent people need to live by certain ideas *given* a certain conception of the agent.

This necessary limitation of the method is not necessarily a problem for it. We can distinguish more generic from more socio-historically local conceptions of the agent, and insofar as we want to tailor a pragmatic genealogy to our particular socio-historical situation, there is nothing problematic about starting with needs human beings have on a generic conception of the agent, and then gradually factoring in the richer, more demanding needs they have on a conception of them that is more peculiarly ours. We must only be aware that this is what we are doing.

Williams's genealogy of liberty, one of the supreme political values in liberal democracies, offers a good example of how a pragmatic genealogy can help us confront the contestation of needs in a way that turns precisely on how parochial its underlying conception of the agent is. In 'From Freedom to Liberty: The Construction of a Political Value', Williams aims to explain not the generic and pre-political value of freedom, but the local and political value of liberty 'as a value *for us* in *our* world' (2005c, 75, emphasis mine). Why, Williams asks, is this idea of liberty so important to us? Why does it play such a central role in *our* political lives?

One response is that liberty and the liberal order are so important to us because human beings have a *need for freedom*. On this view, liberalism is easily justified, because the need for it is built right into the conception of the agent.

But the glaring weakness in this line of argument is that while human beings may have a need for freedom in some sense, they hardly have it *in a sense strong enough* for it to vindicate the liberal order against its historical rivals. It is only if the need for freedom is understood in a particularly demanding sense—as what Williams labels the need for 'autonomy' (2005i, 8)—that it will deliver what this attempt to justify liberalism expects of it.

The flaw in this justification of the value of liberty is thus that the conception of the agent on which human beings have such a demanding form of the need for freedom is itself a *liberal* conception: a recent arrival on the historical scene which fits the liberal order because it emerged alongside it, out of much the same historical forces. And to justify the liberal order in terms of a coeval conception of the agent that only liberals accept is mere self-congratulation. It takes us no further at all in authenticating liberalism as, in any independent sense, the right political arrangement.[38] The argument is not so much wrong as too *internal*: it spells out the value of liberty 'from here'. But once we acknowledge that 'here' is just one place among others—a fact that we historically self-conscious moderns are bound to bump into—we also need an *external* answer to our question to achieve *reasonable* as opposed to blind confidence in the value of liberty. Achieving reasonable confidence requires achieving a vindicatory reflective understanding of liberty as a value: an understanding, among other things, of why we have it, what needs it answers to, and whether it is right for us given our circumstances. This is why, for Williams, the conception of liberty 'we need for ourselves' must be 'historically self-conscious' if it is to be 'suitable to a modern society' (2005c, 75). It must *include a reflective understanding* of the basic concerns to which a more generic notion of freedom answers, and of *why* the socio-historical elaboration of the notion of freedom *we* happen to have is adequate to *our* socio-historical elaboration of those basic concerns. Precisely because we can consider our conceptual practices from the outside and compare them with real and possible alternatives—precisely because we are aware of their contingency—we need to say more about why we have the ones we happen to have, and to what extent they are right *for us*, if we are to sustain reasonable confidence in them. In other words, the understanding of freedom that *we* need includes the kind of understanding yielded by pragmatic genealogy.

Williams makes a step towards such an understanding by sketching a vindicatory genealogy of liberty which starts out from the need for what he calls 'primitive freedom'—the pre-political notion of an individual's freedom from constraint by other individuals in trying to satisfy his or her desires. This is a need that is not contested in the way that the demanding need for autonomy is. As Williams

---

[38] See Williams (2005d, 74; 2005e, 20–3; 2005h, 39; 2005i, 7–9; 2005j, 133).

remarks, we can hardly make sense of a conception of human agents on which they have no need for even the most primitive form of freedom:

> Why should human beings in general be concerned with some value of that form? I do not know that I can answer that question, beyond suggesting a set of questions to put in its place: What view would one have to take of one's desires and projects and other values if there were never even a question of its being something to be resented and resisted if others aimed to frustrate them? What view would one have to take of those others, in particular of a political authority, for that question never to arise?   (2005c, 93)

To begin from a conception of agents on which they care about being unobstructed by others in doing what they want is not to project one's liberal concerns into the starting point; it is to begin from a conception that is, at the very least, far more widely shared than that, and likely even without alternative.

From this starting point, Williams argues that in pursuing their primitive freedom, individuals will impinge on each other's freedom spheres, and one individual's desire satisfaction will be another individual's coercion. A basic problem emerges: where does one freedom sphere end and the other begin? Disagreement over this generates violence, instability, and chaos. This gives rise to the need for a public conflict-resolver, an allocator of freedom spheres. But if this allocator of freedom spheres is not to replace private with public coercion (in which case it remains a mere example of successful banditry that reproduces the problem of coercion at a higher level), there needs to be a distinction between *legitimate* and *illegitimate* uses of public power. Consequently, there is a need for legitimating concepts that permit this distinction. This need is scalable and context-sensitive, and it will become more pressing the closer the situation comes to that in which some public power uniquely commands the means of coercion—that is, the closer it comes to the ideal type of state power. But wherever this need is manifest, the required legitimating concepts will have to be fleshed out in terms of a *legitimation story*, which, drawing, for instance, on religious or transcendent sources of authority, will explain to each citizen why public power can be used to coerce certain people in certain ways. The function of these legitimating concepts and the legitimation story they articulate is thus to secure a political form of freedom under public power by putting normative constraints on that power and justifying certain exercises of it to each citizen.

The basic political problem highlighted by Williams's genealogy is that we need *some* legitimating concepts enabling a distinction between good and bad government. But these needs cannot by themselves determine *which* concepts these will be—whether the legitimation stories will draw on the idea of liberty, for example, or on theological or transcendent sources of legitimacy. Under conditions of modernity, moreover, truthful inquiry and historical self-consciousness have

eroded many of the myths, narratives, and Whiggish histories that formed the stuff of past legitimation stories, leaving us with less material for our legitimation stories; and once these sources of legitimation have fallen away, there is a stronger presumption in favour of citizens' freedom to do what they decidedly want. This helps explain the special importance of liberty under conditions of modernity. We are more concerned with liberty than past societies because 'we start, in a sense, with less' (2005c, 95)—in particular, less by which to justify restricting liberty. Our stronger presumption in favour of the liberty of the individual reflects the fact that fewer sources of legitimation are available to us, thus barring us from justifying exercises of power that could formerly have seemed legitimate.

This genealogy of liberty helps explain our special concern with liberty, but also shows that we are *rightly* more concerned with liberty by presenting our heightened concern with liberty as an expression of truthfulness. As Williams writes, it

> connects our construction of liberty, and the value we give it under that construction, with the condition of modernity, but it offers more than the consideration (which is in itself a perfectly sound consideration) that this is our condition. It connects our ideas of liberty with a universal truth, that everywhere legitimacy requires more than mere coercion, and it adds to this the conviction that under the conditions of modernity, whatever else may be worse, we at any rate have a better grasp on the truth.   (2005c, 95–6)

Williams's genealogy is vindicatory, but it does not ground the need for liberalism in a demanding need for liberty that is itself just as contested as liberalism itself. Rather, it presents our special concern with liberty as a local manifestation of a near-universal predicament, a manifestation reflecting practical pressures that are distinctive of our situation; and it simultaneously presents the fact that we brought ourselves into this situation *as an achievement*. Pragmatic genealogies that, like this one, are tailored to a specific situation can yield a vindicatory understanding of how that situation and the conceptions that are peculiar to it relate to both past and possible alternatives.

To conclude, then, pragmatic genealogies can help us navigate contestations of needs by giving us a model or plan of how our ways of thinking relate to both local and less local needs, thereby allowing us to move beyond the simple contrast between 'what is necessary and what is the product merely of our own contingent arrangements' (Skinner 1969, 53). They enable us to distinguish between *different ways* in which things can be necessary: the ways in which things are necessary *for us* for different values of 'us'. This is what Williams sees as the chief contribution that genealogy can make; that it

> can help with the business, which is quite certainly a philosophical business, of distinguishing between different ways in which various of our ideas and

procedures can seem to be such that we cannot get beyond them, that there is no conceivable alternative.... Wittgenstein influentially and correctly insisted that there was an end to justifications, that at various points we run into the fact that 'this is the way we go on'. But...it makes a great difference who 'we' are supposed to be, and it may mean different groups in different philosophical connections. It may mean maximally...any creature that you and I could conceive of understanding. Or it may mean any human beings, and here universal conditions of human life, including very general psychological capacities, may be relevant. Or it may mean just those with whom you and I share much more, such as outlooks typical of modernity.   (2006d, 195–6)

As citizens of modern liberal democracies, we may have needs, such as a need for political freedom and self-determination, that human beings in different socio-historical situations did not have in this form or to that degree. In trying to provide a comprehensive view of our conceptual practices as rooted in a complex historical accumulation of both generic and socio-historically local needs, many of which depend on or derive from each other, pragmatic genealogy attempts to do justice to these differences. Its aim is not to arrive at something incontestable, but to provide us with a deeper and more nuanced understanding of what we are contesting.

10

Ideas Worth Having

Although any human community needs some ideas to live by, it is much less clear that there is any one set of ideas this should be. Under conditions of modernity, where societies are pluralistic and historically self-conscious, we are inescapably aware that there are real alternatives to the ideas we happen to have. This forces a question on us: how confident should we be in the ideas we live by?

We can now see that pragmatic genealogies provide one kind of answer to this question: they show to what extent the ideas we live by are *ideas worth having*. The pragmatic genealogies we have considered in this book do this by relating ideas we are less confident we need to what we are more confident we need. It is because many ideas have or lack a point for us in certain contexts without being *understood* to have or lack a point for us that pragmatic genealogies can have a vindicatory or subversive effect. All that is required for ideas to have a point for us is that their application in a given context stand in an instrumental relation to our needs. The reflective activity of representing that relation to ourselves is then a further step—the step we take in raising Pragmatic Questions and answering them with pragmatic genealogies.

10.1 Grounding Socratic in Pragmatic Inquiry

Pragmatic Questions are not just an underexplored alternative to Socratic 'What is X?' questions. They are something that Socratic Questions can be grounded in and guided by. It is all very well to pursue Socratic Questions in order to bring into focus the guidance that the ideas we already live by can provide—we can ask what a *good life* is, or what a *democratic society* is; by analysing these notions, we achieve greater clarity about the reasons, norms, and proprieties they encode and the demands they make on us. But the clearer we become about what our ideas demand of us, the more pressing the question becomes of what *authority* these ideas have over us. Why should we heed the demands that our ideas make on us? Why, in particular, should we heed the demands of just *these* ideas *thus* conceived as opposed to those of alternative ideas or conceptions we might also live by? What does it *do* for us to acknowledge the authority of ideas that demand these things of us?

These are questions that Socratic inquiry by itself cannot answer, because it must always beg the question, once the analysed concept lies before us like a

patient etherized upon a table, of why it is this concept rather than another that we should live by. The answer to a 'What is X?' question cannot be the last word, because the authority of any answer to that question needs to be assessed in light of the merit of accepting it as authoritative. This is where Pragmatic Questions come in to fill a need. They allow us to disengage ourselves from a given idea and take a sideways view of it in order to assess its right to have a hold on us according to whether this helps us to live. This kind of inquiry allows us to assess what ideas *demand* of us in light of what they *do* for us, and this just is to assess answers to Socratic Questions in light of answers to Pragmatic Questions. Socratic inquiry can help us understand our ideas better, but it cannot tell us whether we are drifting away from ideas we have reason to cultivate, or blindly adhering to ideas we have reason to abandon. That judgement requires the kind of self-understanding yielded by pragmatic genealogies.

10.2 Pursuing Philosophy as a Humanistic Discipline

While good pragmatic genealogies lead to better self-understanding, there is also an important sense in which it is the other way around: good self-understanding leads to better pragmatic genealogies. How one sets up one's genealogy, what facts, needs, and problems one sets out from, are choices grounded in psychological, social, and historical understanding. One starts somewhere according to one's best knowledge of what kinds of creatures we most basically are, and adds further needs according to one's best knowledge of what we more particularly came to be. The primary aim of a pragmatic genealogy is to bring out what a concept does for us when it functions well, and to understand this, one needs to bring into view the broad array of contexts, practices, and institutions in which it is put to work—the rough ground with which it makes contact when, after having been held up for philosophical inspection, it is lowered back into human affairs.

It is therefore only at a superficial level that the present interpretation could be said to sever the connection between genealogy and history; at a deeper level, the connection is very much there, as it is history, along with the other human sciences, that provides the understanding that guides and informs the selection of what should go into the dynamic model of a pragmatic genealogy. Pragmatic genealogy does not compete with more regularly historiographical or Foucauldian forms of genealogy. On the contrary, it feeds off them, and should be informed by them. Otherwise, we run the risk of overgeneralizing from our own experience and, more generally, of operating with an unrealistic view of human affairs. As Lewis Namier remarked, history gives us an intuitive understanding of how things do *not* happen.[1]

---

[1] See Stern (1956, 30).

Yet the integration with neighbouring disciplines does not just function as a check on genealogists' prejudices. What pragmatic genealogists most basically seek to do is to render transparent the relation of our concepts to our needs; but the needs that function as input to genealogies' dynamic models have to come from somewhere, as does the largely tacit conception of human capacities and the circumstances in which they are exercised that we draw on in moving from one stage of the dynamic model to the next. If the practical value of concepts can only be assessed against the background of a nuanced understanding of the world in which they are put to work, then philosophy must engage with and incorporate the findings of other disciplines in the humanities and the social sciences. The guiding concerns will still recognizably be those of philosophy; but in pursuing its own concerns, philosophy need not entirely abandon those of the humanities and social sciences. Philosophy 'cannot be too pure if it really wants to do what it sets out to do' (2002, 39), Williams insisted, and 'if there is to be a philosophical way of doing better respect to the complexity of these value concepts and their relations to a wider background, it will not be one which totally leaves behind the interests of the social sciences, but rather one which cooperates with them' (2006g, 158).[2]

What other disciplines can tell us about concept-users and their wider circumstances does not just act as input to pragmatic genealogy, moreover. It also offers an external validation of its output. The genealogies we considered give us reasons to expect certain conceptual practices to be extremely widespread across many human cultures and epochs. Consequently, observation and comparison of human cultures and epochs might prove them wrong or incomplete in this respect. Should it turn out that ideas of property, knowledge, or truthfulness are only very recent arrivals on the scene of world history, or that they are far more parochial than these genealogies suggest, this will show the assumptions that went into constructing the models to be, at least by themselves, inadequate. A genealogy's appeal will be directly proportional to its ability to explain and predict the shape of our conceptual practices as characterized by the various disciplines that are in the business of describing how human beings actually think, speak, and act.

Hence, the facts about the needs, capacities, and circumstances of concept-users that go into and come out of pragmatic genealogies are drawn from and validated by the varieties of empirical observation that form the province of the human sciences. What the dynamic models of pragmatic genealogy offer are *receptacles* for the incorporation of social, psychological, and historical understanding into philosophy.

One might thus say that pragmatic genealogy provides a concrete proposal for how to pursue philosophy as a humanistic discipline. The method is designed to

---

[2] See Moran (2016) for further discussion of Williams's conception of the 'impurity' of philosophy.

help us negotiate the border between philosophy and the human sciences. Its dynamic models exemplify a way in which philosophy can be practised as a humanistic discipline, i.e. in close integration with neighbouring human sciences, without surrendering its disciplinary identity. In their abstraction, their emphases, and the types of insight they yield, these dynamic models are recognizably the tools of philosophers rather than of inquirers specialized in the nit-and-grit of human affairs. Considered as regular psychology, sociology, or historiography, the pragmatic genealogies of Hume, Nietzsche, Craig, Williams, and Fricker are odd creatures indeed. But their contours make perfect sense when considered as answers to philosophical concerns about the naturalistic credentials or the point and value of certain concepts. Here, idealizing and potentially distorting narrative devices like the state of nature serve a point, because the concerns of philosophy are not best served by a 'Laplacean' genealogy capturing every little detail in the meandering history of our practices. Here less is more, for it distils into a perspicuous narrative the essential dynamics we need to grasp in order for our suspicions to be awakened or assuaged. At the same time, this condensed narrative is still a long way from a priori philosophical reflection as exemplified by transcendental arguments. It would not satisfy a purist definition of philosophy as reflection that remains independent from how the world contingently is, because the genealogical narrative aims to embody a form of philosophical reflection that is suitably sensitive to the contingent dependencies of our concepts.

True to the spirit of its subject matter, the present book has sought to uncover the point of giving pragmatic genealogies, the practical and theoretical needs answered thereby, and the method's elaboration in response to the specific concerns of Hume, Nietzsche, Craig, Williams, and Fricker. It has emerged that there is a pragmatic genealogical tradition with which contemporary analytic philosophy should find it comparatively easy to identify. And the tradition lives on. Philip Pettit's pragmatic genealogy of moral desirability and responsibility in *The Birth of Ethics* (2018) neatly aligns itself with the trajectory I have been tracing, not least because it explicitly labels its state-of-nature model with an anagram for 'nowhere' to pre-empt the misunderstanding that it depicts early hominins on Pleistocene plains. It has been a guiding theme of the present book that pragmatic genealogies are best interpreted as depictions not of our actual present or distant past, but of nowhere in particular. They are models serving to reveal whether and when our ideas are worth having.

The other guiding theme, finally, has been that while pragmatism, naturalism, and genealogy can each invite the charge of taking an overly reductive view of human concerns when taken singly, their fusion in pragmatic genealogy yields a framework capable of vindicating both the insight that our concepts grow out of our needs *and* the insight that need-satisfaction is often not our concern. Often, our needs shape the space of reasons the way our eyes shape the field of vision: they render us sensitive to certain aspects of the world without themselves

showing up in it. Pragmatic genealogy allows us to represent to ourselves and others these relations between our conceptual practices and our needs. What material a genealogy should draw on will depend on what commands allegiance among its addressees, just as the point of telling it will depend on the concerns, expectations, and needs in response to which it is offered. The narrative device of pragmatic genealogy is itself a tool, imbued with a point by the need of its addressees to understand themselves better, and by its ability to show them whether the ideas they live by serve needs they endorse. For those whose ideas hold no more surprises in this regard, that tool is pointless. But not for us.

Bibliography

Alcoff, Linda Martín. 2010. 'Epistemic Identities'. *Episteme* 7 (2): 128–37.

Anderson, Elizabeth. 2001. 'Unstrapping the Straitjacket of "Preference": A Comment on Amartya Sen's Contributions to Philosophy and Economics'. *Economics and Philosophy* 17 (1): 21–38.

Anderson, Elizabeth. 2012. 'Epistemic Justice as a Virtue of Social Institutions'. *Social Epistemology* 26 (2): 163–73.

Angehrn, Emil. 2007. *Die Frage nach dem Ursprung: Philosophie zwischen Ursprungsdenken und Ursprungskritik*. Munich: Fink.

Anghie, Antony. 2007. *Imperialism, Sovereignty and the Making of International Law*. Cambridge: Cambridge University Press.

Appiah, Kwame Anthony. 2017. *As If: Idealization and Ideals*. Cambridge, MA: Harvard University Press.

Arendt, Hannah. 1958. *The Human Condition*. Chicago, IL: University of Chicago Press.

Austin, John Langshaw. 1961. 'A Plea for Excuses'. In *Philosophical Papers*. Edited by J. O. Urmson and G. J. Warnock, 123–52. Oxford: Clarendon Press.

Austin, John Langshaw. 1962. *Sense and Sensibilia*. Edited by G. J. Warnock. Oxford: Oxford University Press.

Baier, Annette. 2010. *The Cautious Jealous Virtue: Hume on Justice*. Cambridge, MA: Harvard University Press.

Baillie, James. 2000. *Hume on Morality*. London: Routledge.

Baker, Kenneth. 2002. 'Bernard Williams: Carrying the Torch for Truth—Philosopher Calculates the Moral Cost of Rejecting the Concept'. *San Francisco Chronicle*, 22 September.

Baldwin, Jason. 2004. 'Hume's Knave and the Interests of Justice'. *Journal of the History of Philosophy* 42 (3): 277–96.

Barnes, Barry. 1995. *The Elements of Social Theory*. Cambridge: Cambridge University Press.

Basu, Rima. 2019. 'The Wrongs of Racist Beliefs'. *Philosophical Studies* 176 (9): 2497–515.

Beam, Craig. 1996. 'Hume and Nietzsche: Naturalists, Ethicists, Anti-Christians'. *Hume Studies* 22 (2): 299–324.

Beam, Craig. 2001. 'Ethical Affinities: Nietzsche in the Tradition of Hume'. *International Studies in Philosophy* 33 (3): 87–98.

Beebe, James R. 2012. 'Social Functions of Knowledge Attributions'. In *Knowledge Ascriptions*. Edited by Jessica Brown and Mikkel Gerken, 220–42. Oxford: Oxford University Press.

Bentham, Jeremy. 1988. *A Fragment on Government*. Edited by James Henderson Burns and Herbert Lionel Adolphus Hart. Cambridge: Cambridge University Press.

Berenstain, Nora. 2016. 'Epistemic Exploitation'. *Ergo: An Open Access Journal of Philosophy* 3 (22): 569–90.

Berlin, Isaiah. 1996. 'The Sense of Reality'. In *The Sense of Reality: Studies in Ideas and their History*. Edited by Henry Hardy, 1–39. New York: Farrar, Straus and Giroux.

Berlin, Isaiah. 1997. *The Proper Study of Mankind*. Edited by Henry Hardy and Roger Hausheer. London: Chatto and Windus.

Berlin, Isaiah. 2002. *Liberty*. Edited by Henry Hardy. Oxford: Oxford University Press.

Berlin, Isaiah. 2013. 'Vico and Herder'. In *Three Critics of the Enlightenment: Vico, Hamann, Herder*. Edited by Henry Hardy, 1–300. Princeton, NJ: Princeton University Press.

Berlin, Isaiah. 2014. *Personal Impressions*. Edited by Henry Hardy. Princeton, NJ: Princeton University Press.

Bevir, Mark. 2008. 'What Is Genealogy?'. *Journal of the Philosophy of History* 2 (3): 263–75.

Blackburn, Simon. 1998. *Ruling Passions*. Oxford: Oxford University Press.

Blackburn, Simon. 2002a. 'Quasi-Realism in Moral Philosophy: An Interview with Simon Blackburn'. *Ethic@* 2 (1): 101–14.

Blackburn, Simon. 2002b. 'Where the Tiger Went'. *Times Literary Supplement*, 18 October, 6–7.

Blackburn, Simon. 2005. *Truth: A Guide for the Perplexed*. London: Allen Lane.

Blackburn, Simon. 2008. *How to Read Hume*. London: Granta.

Blackburn, Simon. 2010. 'The Majesty of Reason'. In *Practical Tortoise Raising and Other Philosophical Essays*, 283–307. Oxford: Oxford University Press.

Blackburn, Simon. 2013a. 'Pragmatism in Philosophy: The Hidden Alternative'. *Philosophic Exchange* 41 (1): 2–13.

Blackburn, Simon. 2013b. 'Pragmatism: All or Some?'. In *Expressivism, Pragmatism and Representationalism*. Edited by Huw Price, 67–84. Cambridge: Cambridge University Press.

Blackburn, Simon. 2016. *The Oxford Dictionary of Philosophy*. 3rd ed. Oxford: Oxford University Press.

Blackburn, Simon. 2017. 'Pragmatism: All or Some or All and Some?'. In *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Edited by Cheryl Misak and Huw Price, 61–74. Oxford: Oxford University Press.

Blackburn, Simon. 2019. 'Lonely in Littlemore: Confidence in Ethics and the Limits of Philosophy'. In *Ethics Beyond the Limits: New Essays on Bernard Williams' Ethics and the Limits of Philosophy*. Edited by Sophie Grace Chappell and Marcel van Ackeren, 27–36. London: Routledge.

Blue, Daniel. 2016. *The Making of Friedrich Nietzsche: The Quest for Identity, 1844–1869*. Cambridge: Cambridge University Press.

Boghossian, Paul. 2003. 'Blind Reasoning'. *Aristotelian Society Supplementary Volume* 77 (1): 225–48.

Boghossian, Paul. 2006. *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Oxford University Press.

Boorse, Christopher. 1976. 'Wright on Functions'. *Philosophical Review* 85 (1): 70–86.

Boyd, Richard. 1991. 'Anti-Foundationalism and the Enthusiasm for Natural Kinds'. *Philosophical Studies* 61: 127–48.

Brandhorst, Mario. 2021. *Ethischer Naturalismus: Ein Plädoyer*. Stuttgart: Metzler.

Brandom, Robert. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Brandom, Robert. 2002. 'When Philosophy Paints Its Blue on Gray: Irony and the Pragmatist Enlightenment'. *Boundary* 29 (2): 1–28.

Brandom, Robert. 2004. 'The Pragmatist Enlightenment (and Its Problematic Semantics)'. *European Journal of Philosophy* 12 (1): 1–16.

Brandom, Robert. 2011. 'Vocabularies of Pragmatism: Synthesizing Naturalism and Historicism'. In *Perspectives on Pragmatism: Classical, Recent, and Contemporary*, 116–57. Cambridge, MA: Harvard University Press.

Brandom, Robert. 2015. 'Den Abgrund reflektieren: Vernunft, Genealogie und die Hermeneutik des Edelmuts'. *West End: Neue Zeitschrift für Sozialforschung* 1: 3–26.

Brandom, Robert. 2019a. 'Some Strands of Wittgenstein's Normative Pragmatism, and Some Strains of His Semantic Nihilism'. *Disputatio* 8 (9): 1–29.

Brandom, Robert. 2019b. *A Spirit of Trust: A Reading of Hegel's Phenomenology*. Cambridge, MA: Harvard University Press.

Brassier, Ray. 2016. 'Dialectics Between Suspicion and Trust'. *Stasis* 4 (2): 98–113.

Brigandt, Ingo. 2010. 'The Epistemic Goal of a Concept: Accounting for the Rationality of Semantic Change and Variation'. *Synthese* 177 (1): 19–40.

Brigandt, Ingo, and Esther Rosario. 2020. 'Strategic Conceptual Engineering for Epistemic and Social Aims'. In *Conceptual Engineering and Conceptual Ethics*. Edited by Alexis Burgess, Herman Cappelen, and David Plunkett, 100–24. Oxford: Oxford University Press.

Brobjer, Thomas H. 2008a. *Nietzsche and the English: The Influence of British and American Thinking on His Philosophy*, JHP Books Series. Amherst, NY: Humanity Books.

Brobjer, Thomas H. 2008b. *Nietzsche's Philosophical Context: An Intellectual Biography*, International Nietzsche Studies. Champaign, IL: University of Illinois Press.

Brock, Gillian, ed. 1998. *Necessary Goods: Our Responsibilities to Meet Others' Needs*. Oxford: Rowman and Littlefield.

Brock, Gillian, and David Miller. 2019. 'Needs in Moral and Political Philosophy'. In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. https://plato.stanford.edu/entries/needs.

Broisson, Ivan. 2006. 'Ressentiment und "Wille zur Macht": Nietzsche und Hume über Moral- und Religionskritik'. In *Friedrich Nietzsche: Zwischen Musik, Philosophie und Ressentiment*. Edited by Volker Gerhardt and Renate Reschke, 117–28. Berlin: Akademie Verlag.

Brun, Georg. 2016. 'Explication as a Method of Conceptual Re-engineering'. *Erkenntnis* 81 (6): 1211–41.

Brun, Georg. 2020. 'Conceptual Re-engineering: From Explication to Reflective Equilibrium'. *Synthese* 197 (3): 925–54.

Brunner, Otto, Werner Conze, and Reinhart Koselleck, eds. 1972–97. *Geschichtliche Grundbegriffe: Historisches Lexikon zur politisch-sozialen Sprache in Deutschland*. Stuttgart: Klett-Cotta.

Buckner, Cameron. 2013. 'Morgan's Canon, Meet Hume's Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons'. *Biology and Philosophy* 28 (5): 853–71.

Burgess, Alexis, Herman Cappelen, and David Plunkett, eds. 2020. *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press.

Byrd, B. Sharon, and Joachim Hruschka. 2010. *Kant's Doctrine of Right: A Commentary*. Cambridge: Cambridge University Press.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

Carey, Susan. 2009. *The Origin of Concepts*. Oxford: Oxford University Press.

Carnap, Rudolf. 1950. *Logical Foundations of Probability*. Chicago, IL: University of Chicago Press.

Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

Chalmers, David J. 2011. 'Verbal Disputes'. *Philosophical Review* 120 (4): 515–66.

Clark, Maudemarie. 1990. *Nietzsche on Truth and Philosophy*. Cambridge: Cambridge University Press.

Clark, Maudemarie. 1998. 'Introduction'. In *On the Genealogy of Morality*, vii–xxxv. Indianapolis, IN: Hackett.

Clark, Maudemarie. 2015a. 'Nietzsche's Immoralism and the Concept of Morality'. In *Nietzsche on Ethics and Politics*, 23–40. Oxford: Oxford University Press.

Clark, Maudemarie. 2015b. 'On the Rejection of Morality: Bernard Williams's Debt to Nietzsche'. In *Nietzsche on Ethics and Politics*, 41–61. Oxford: Oxford University Press.

Coady, David. 2010. 'Two Concepts of Epistemic Injustice'. *Episteme* 7 (2): 101–13.

Cohen, Morris R., and Ernest Nagel. 1934. *Introduction to Logic and Scientific Method*. New York: Harcourt, Brace and Company.

Cohon, Rachel. 1997. 'Hume's Difficulty with the Virtue of Honesty'. *Hume Studies* 23 (1): 91–112.

Cohon, Rachel. 2006. 'Hume's Artificial and Natural Virtues'. In *The Blackwell Guide to Hume's Treatise*. Edited by Saul Traiger, 256–75. Oxford: Blackwell.

Cohon, Rachel. 2008. *Hume's Morality: Feeling and Fabrication*. New York: Oxford University Press.

Cohon, Rachel. 2010. 'Hume's Moral Philosophy'. In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. https://plato.stanford.edu/entries/hume-moral.

Cottingham, John. 2009. 'The Good Life and the "Radical Contingency of the Ethical"'. In *Reading Bernard Williams*. Edited by Daniel Callcut, 24–42. London: Routledge.

Craig, Edward. 1986. 'The Practical Explication of Knowledge'. *Proceedings of the Aristotelian Society* 87: 211–26.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon Press.

Craig, Edward. 1993. *Was wir wissen können: Pragmatische Untersuchungen zum Wissensbegriff. Wittgenstein-Vorlesungen der Universität Bayreuth*. Edited by Wilhelm Vossenkuhl. Frankfurt am Main: Suhrkamp.

Craig, Edward. 2007. 'Genealogies and the State of Nature'. In *Bernard Williams*. Edited by Alan Thomas, 181–200. Cambridge: Cambridge University Press.

Crawley, Richard. 1998. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War*. Edited by Robert B. Strassler. New York: Simon and Schuster.

Crouch, Margaret A. 1993. 'A "Limited" Defense of the Genetic Fallacy'. *Metaphilosophy* 24 (3): 227–40.

Cueni, Damian. 2020. 'Public Law Analogies in International Legal Theory'. Doctoral Thesis, Rechtswissenschaftliches Institut, University of Zurich.

Cueni, Damian, and Matthieu Queloz. 2021. 'Whence the Demand for Ethical Theory?'. *American Philosophical Quarterly* 58 (2): 135–45.

Cueni, Damian, and Matthieu Queloz. Manuscript. 'The Normative Significance of Critical Histories for International Law'.

Cullison, Andrew. 2010. 'On the Nature of Testimony'. *Episteme* 7 (2): 114–27.

Cummins, Robert C. 1975. 'Functional Analysis'. *Journal of Philosophy* 72: 741–64.

Darwall, Stephen. 1993. 'Motive and Obligation in Hume's Ethics'. *Noûs* 27 (4): 415–48.

Darwall, Stephen. 1995. *The British Moralists and the Internal 'Ought': 1640–1740*. Cambridge: Cambridge University Press.

Daston, Lorraine, and Peter Galison. 1992. 'The Image of Objectivity'. *Representations* 40: 81–128.

Davidson, Donald. 1990. 'The Structure and Content of Truth'. *Journal of Philosophy* 87 (6): 279–328.

Davis, Emmalon. 2016. 'Typecasts, Tokens, and Spokespersons: A Case for Credibility Excess as Testimonial Injustice'. *Hypatia* 31 (3): 485–501.

Descartes, René. 1996. *Œuvres de Descartes*. Edited by Charles Adam and Paul Tannery. Paris: Vrin.

Dogramaci, Sinan. 2012. 'Reverse Engineering Epistemic Evaluations'. *Philosophy and Phenomenological Research* 84 (3): 513–30.

Dotson, Kristie. 2011. 'Tracking Epistemic Violence, Tracking Practices of Silencing'. *Hypatia* 26 (2): 236–57.

Driscoll, Catherine. 2015. 'Neither Adaptive Thinking nor Reverse Engineering: Methods in the Evolutionary Social Sciences'. *Biology and Philosophy* 30 (1): 59–75.

Drossbach, Maximilian. 1884. *Über die scheinbaren und die wirklichen Ursachen des Geschehens in der Welt*. Halle/Saale: Pfeffer.

Dummett, Michael. 1993. *Origins of Analytic Philosophy*. London: Duckworth.

Dummett, Michael. 1996. *The Seas of Language*. Oxford: Clarendon Press.

Dummett, Michael. 2004. *Truth and the Past*. New York: Columbia University Press.

Dupré, John. 1998. 'Against Reductionist Explanations of Human Behaviour'. *Aristotelian Society Supplementary Volume* 72 (1): 153–72.

Dupré, John. 2001. *Human Nature and the Limits of Science*. Oxford: Oxford University Press.

Dutilh Novaes, Catarina. 2015. 'Conceptual Genealogy for Analytic Philosophy'. In *Beyond the Analytic–Continental Divide*. Edited by Jeffrey A. Bell, Andrew Culrofello, and Paul M. Livingston, 75–110. London: Routledge.

Dutilh Novaes, Catarina. 2020. 'Carnapian Explication and Ameliorative Analysis: A Systematic Comparison'. *Synthese* 197 (3): 1001–34.

Eklund, Matti. 2007. 'Meaning-Constitutivity'. *Inquiry* 50 (6): 559–74.

Elgin, Catherine Z. 2005. 'Williams on Truthfulness'. *The Philosophical Quarterly* 55 (219): 343–52.

Elgin, Catherine Z. 2007. 'Understanding and the Facts'. *Philosophical Studies* 132 (1): 33–42.

Elgin, Catherine Z. 2020. 'The Mark of a Good Informant'. *Acta Analytica* 35 (3): 319–331.

Emden, Christian J. 2014. *Nietzsche's Naturalism: Philosophy and the Life Sciences in the Nineteenth Century*. Cambridge: Cambridge University Press.

Epstein, Brian. 2010. 'History and the Critique of Social Concepts'. *Philosophy of the Social Sciences* 40 (1): 3–29.

Evnine, Simon. 1993. 'Hume, Conjectural History, and the Uniformity of Human Nature'. *Journal of the History of Philosophy* 31 (4): 589–606.

Faguet, Émile. 1891. *Politiques et moralistes du XIXe siècle: première série*. Paris: Lecène, Oudin et cie.

Faulkner, Paul. 2007. 'A Genealogy of Trust'. *Episteme* 4 (3): 305–21.

Ferguson, Adam. 1966. *An Essay on the History of Civil Society*. Edinburgh: University of Edinburgh Press.

Finnis, John. 2011. 'Bernard Williams on Truth's Values'. In *Reason in Action: Collected Essays: Volume I*, 92–103. Oxford: Oxford University Press.

Fischer, Kuno. 1869. *Kant's Vernunftkritik und deren Entstehung*. Heidelberg: Bassermann.

Fleischacker, Samuel. 2004. 'Review of Bernard Williams, Truth and Truthfulness: An Essay in Genealogy'. *Ethics* 114 (2): 380–5.

Fodor, Jerry A. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.

Fodor, Jerry A. 2004. 'Having Concepts: A Brief Refutation of the Twentieth Century'. *Mind and Language* 19 (1): 29–47.

Forrester, Katrina. 2019. *In the Shadow of Justice: Postwar Liberalism and the Remaking of Political Philosophy*. Princeton, NJ: Princeton University Press.

Foucault, Michel. 1971. 'Nietzsche, la généalogie, l'histoire'. In *Hommage à Jean Hyppolite*. Edited by S. Bachelard, 145–72. Paris: Presses Universitaires de France.

Foucault, Michel. 1975. *Surveiller et punir: Naissance de la prison*. Paris: Gallimard.

Foucault, Michel. 1994. 'La vérité et les formes juridiques'. In *Dits et Écrits 1954–1988*. Vol. II (1970–5), 538–646. Paris: Gallimard.

Fraser, Nancy. 1981. 'Foucault on Modern Power: Empirical Insights and Normative Confusions'. In *Unruly Practices: Power, Discourse and Gender in Contemporary Social Theory*. Edited by Nancy Fraser, 17–34. Minneapolis, MN: University of Minnesota Press.

Frege, Gottlob. 1893. *Grundgesetze der Arithmetik*. Vol. 1. Jena: Hermann Pohle.

Fricker, Elizabeth. 2015. 'Know First, Tell Later: The Truth about Craig on Knowledge'. In *Epistemic Evaluation: Purposeful Epistemology*. Edited by John Greco and David Henderson, 46–86. Oxford: Oxford University Press.

Fricker, Miranda. 1998. 'Rational Authority and Social Power: Towards a Truly Social Epistemology'. *Proceedings of the Aristotelian Society* 98 (2): 159–77.

Fricker, Miranda. 2000. 'Confidence and Irony'. In *Morality, Reflection, and Ideology*. Edited by Edward Harcourt, 87–112. Oxford: Oxford University Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Fricker, Miranda. 2008. 'Scepticism and the Genealogy of Knowledge: Situating Epistemology in Time'. *Philosophical Papers* 37 (1): 27–50.

Fricker, Miranda. 2009. 'The Value of Knowledge and the Test of Time'. *Royal Institute of Philosophy Supplement* 64: 121–38.

Fricker, Miranda. 2010a. 'Replies to Alcoff, Goldberg, and Hookway on Epistemic Injustice'. *Episteme* 7 (2): 164–78.

Fricker, Miranda. 2010b. 'Scepticism and the Genealogy of Knowledge: Situating Epistemology in Time'. In *Social Epistemology*. Edited by A. Haddock, A. Millar, and D. Pritchard, 51–68. Oxford: Oxford University Press.

Fricker, Miranda. 2012. 'Group Testimony? The Making of a Collective Good Informant'. *Philosophy and Phenomenological Research* 84 (2): 249–76.

Fricker, Miranda. 2013. 'Epistemic Justice as a Condition of Political Freedom?'. *Synthese* 190 (7): 1317–32.

Fricker, Miranda. 2015. 'Epistemic Contribution as a Central Human Capability'. In *The Equal Society: Essays on Equality in Theory and Practice*. Edited by George Hull, 73–89. Lanham, MD: Lexington Books.

Fricker, Miranda. 2016a. 'Fault and No-Fault Responsibility for Implicit Prejudice: A Space for Epistemic "Agent-Regret"'. In *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*. Edited by Michael S. Brady and Miranda Fricker, 33–50. Oxford: Oxford University Press.

Fricker, Miranda. 2016b. 'What's the Point of Blame? A Paradigm Based Explanation'. *Noûs* 50 (1): 165–83.

Fricker, Miranda. 2017. 'Evolving Concepts of Epistemic Injustice'. In *The Routledge Handbook of Epistemic Injustice*. Edited by Ian James Kidd, José Medina, and Gaile Pohlhaus Jr, 53–60. New York: Routledge.

Fricker, Miranda. Forthcoming. 'Forgiveness: An Ordered Pluralism'. *Australasian Philosophical Review*.

Frigg, Roman, and Stephan Hartmann. 2017. 'Models in Science'. In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. https://plato.stanford.edu/entries/models-science.

Galilei, Galileo. 1954. *Dialogues Concerning Two New Sciences*. New York: Dover Publications.

Gardiner, Georgi. 2015. 'Teleologies and the Methodology of Epistemology'. In *Epistemic Evaluation: Purposeful Epistemology*. Edited by John Greco and David Henderson, 31–45. Oxford: Oxford University Press.

Garrett, Don. 2007. 'The First Motive to Justice: Hume's Circle Argument Squared'. *Hume Studies* 33 (2): 257–88.

Garrett, Don. 2015. *David Hume*. London: Routledge.

Garson, Justin. 2016. *A Critical Overview of Biological Functions*. Dordrecht: Springer.

Garson, Justin. 2019. *What Biological Functions Are and Why They Matter*. Cambridge: Cambridge University Press.

Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.

Gauthier, David. 1992. 'Artificial Virtue and the Sensible Knave'. *Hume Studies* 18 (2): 401–28.

Gay, Peter. 1971. *The Party of Humanity: Essays in the French Enlightenment*. New York: W. W. Norton.

Geach, Peter. 1977. *The Virtues: The Stanton Lectures, 1973–4*. Cambridge: Cambridge University Press.

Gelfert, Axel. 2014. *A Critical Introduction to Testimony*. London: Bloomsbury Academic.

Gemes, Ken. 1992. 'Nietzsche's Critique of Truth'. *Philosophy and Phenomenological Research* 52 (1): 47–65.

Gemes, Ken. 2006. '"We Remain of Necessity Strangers to Ourselves": The Key Message of Nietzsche's Genealogy'. In *Nietzsche's On the Genealogy of Morals: Critical Essays*. Edited by Christa Davis Acampora, 191–208. Lanham, MD: Rowman and Littlefield.

Gerken, Mikkel. 2017. *On Folk Epistemology: How We Think and Talk about Knowledge*. Oxford: Oxford University Press.

Gettier, Edmund. 1963. 'Is Justified True Belief Knowledge?'. *Analysis* 23 (6): 121–3.

Geuss, Raymond. 1999. 'Nietzsche and Genealogy'. In *Morality, Culture and History*, 1–28. Cambridge: Cambridge University Press.

Giere, Ronald N. 1999. *Science Without Laws*. Chicago, IL: University of Chicago Press.

Glock, Hans-Johann. 2006a. 'Concepts: Representations or Abilities?'. In *Content, Consciousness, and Perception: Essays in Contemporary Philosophy of Mind*. Edited by Ezio Di Nucci and Conor McHugh, 36–61. Cambridge: Cambridge Scholars Press.

Glock, Hans-Johann. 2006b. 'Wittgenstein and History'. In *Wittgenstein: The Philosopher and His Works*. Edited by A. Pichler and S. Säätelä, 277–303. Frankfurt am Main: Ontos.

Glock, Hans-Johann. 2008a. 'Analytic Philosophy and History: A Mismatch?'. *Mind* 117 (468): 867–97.

Glock, Hans-Johann. 2008b. *What Is Analytic Philosophy?* Cambridge: Cambridge University Press.

Glock, Hans-Johann. 2009. 'Concepts: Where Subjectivism Goes Wrong'. *Philosophy* 84 (1): 5–29.

Glock, Hans-Johann. 2010. 'Concepts: Between the Subjective and the Objective'. In *Mind, Method, and Morality: Essays in Honour of Anthony Kenny*. Edited by J. Cottingham and P. M. S. Hacker, 306–29. Oxford: Oxford University Press.

Glock, Hans-Johann. 2017a. 'Philosophy and Philosophical Method'. In *A Companion to Wittgenstein*. Edited by Hans-Johann Glock and John Hyman, 231–51. Chichester: Wiley-Blackwell.

Glock, Hans-Johann. 2017b. 'Wittgenstein's Rain in the "Philosophical Desert": Pragmatist Ideas in Post-War Oxford'. In *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Edited by Cheryl Misak and Huw Price, 131–48. Oxford: Oxford University Press.

Goldberg, Sanford. 2010. 'Comments on Miranda Fricker's Epistemic Injustice'. *Episteme* 7 (2): 138–50.

Goldie, Peter. 2009. 'Thick Concepts and Emotion'. In *Reading Bernard Williams*. Edited by Daniel Callcut, 94–109. London: Routledge.

Goudge, Thomas Anderson. 1961. 'The Genetic Fallacy'. *Synthese* 13 (1): 41–8.

Graeber, David. 2011. *Debt: The First 5000 Years*. London: Melville.

Greco, Daniel, and Brian Hedden. 2016. 'Uniqueness and Metaepistemology'. *The Journal of Philosophy* 113 (8): 365–95.

Greco, John, and David Henderson, eds. 2015. *Epistemic Evaluation: Purposeful Epistemology*. Oxford: Oxford University Press.

Greco, Lorenzo. 2007. 'Humean Reflections in the Ethics of Bernard Williams'. *Utilitas* 19: 311–25.

Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grimm, Stephen R. 2015. 'Knowledge, Practical Interests, and Rising Tides'. In *Epistemic Evaluation: Purposeful Epistemology*. Edited by John Greco and David Henderson, 117–37. Oxford: Oxford University Press.

Grimm, Stephen R., Christoph Baumberger, and Sabine Ammon, eds. 2016. *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*. London: Routledge.

Griswold, Charles L. 2018. *Jean-Jacques Rousseau and Adam Smith: A Philosophical Encounter*. London: Routledge.

Gutting, Gary. 2005. *Foucault*. Oxford: Oxford University Press.

Habermas, Jürgen. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by William Rehg. Cambridge: Polity.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Hacking, Ian. 2004. 'Critical Notice'. *Canadian Journal of Philosophy* 34 (1): 137–48.

Hacking, Ian. 2005. 'Truthfulness'. *Common Knowledge* 11 (1): 160–72.

Hall, Edward. 2014. 'Contingency, Confidence, and Liberalism in the Political Thought of Bernard Williams'. *Social Theory and Practice* 40 (4): 545–69.

Hampton, Jean. 1987. *Hobbes and the Social Contract Tradition*. Cambridge: Cambridge University Press.

Hanfling, Oswald. 1985. 'A Structural Account of Knowledge'. *The Monist* 68 (1): 40–57.

Hannon, Michael. 2013. 'The Practical Origins of Epistemic Contextualism'. *Erkenntnis* 78 (4): 899–919.

Hannon, Michael. 2015. 'The Universal Core of Knowledge'. *Synthese* 192 (3): 769–86.

Hannon, Michael. 2019. *What's the Point of Knowledge? A Function-First Epistemology*. Oxford: Oxford University Press.

Hanson, Norwood Russell. 1967. 'The Genetic Fallacy Revisited'. *American Philosophical Quarterly* 4 (2): 101–13.

Harrison, Jonathan. 1981. *Hume's Theory of Justice*. Oxford: Clarendon Press.

Hart, Herbert Lionel Adolphus. 1961. *The Concept of Law*. Oxford: Oxford University Press.

Hartmann, Martin, and Martin Saar. 2004. 'Bernard Williams on Truth and Genealogy'. *European Journal of Philosophy* 12 (3): 386–98.

Haslanger, Sally. 2000. 'Gender and Race: (What) Are They? (What) Do We Want Them to Be?'. *Noûs* 34 (1): 31–55.

Haslanger, Sally. 2012a. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.

Haslanger, Sally. 2012b. 'What Are We Talking About? The Semantics and Politics of Social Kinds'. In *Resisting Reality: Social Construction and Social Critique*, 365–80. New York: Oxford University Press.

Haslanger, Sally. 2014. 'Social Meaning and Philosophical Method'. *Proceedings and Addresses of the American Philosophical Association* 88: 16–37.

Haslanger, Sally. 2019. 'Cognition as a Social Skill'. *Australasian Philosophical Review* 3 (1): 5–25.

Haslanger, Sally. 2020. 'Going On, Not in the Same Way'. In *Conceptual Engineering and Conceptual Ethics*. Edited by Alexis Burgess, Herman Cappelen, and David Plunkett, 230–60. Oxford: Oxford University Press.

Haslanger, Sally. Forthcoming. 'How Not to Change the Subject'. In *Shifting Concepts: The Philosophy and Psychology of Conceptual Variation*. Edited by Teresa Marques and Åsa Wikforss. Oxford: Oxford University Press.

Hazlett, Allan. 2013. *A Luxury of the Understanding: On the Value of True Belief*. Oxford: Oxford University Press.

Heal, Jane. 2013. 'Social Anti-Individualism, Co-Cognitivism, and Second Person Authority'. *Mind* 122 (486): 339–71.

Heal, Jane. 2017. 'Pragmatism and Anscombe on the First Person'. In *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Edited by Cheryl Misak and Huw Price, 117–30. Oxford: Oxford University Press.

Henderson, David. 2011. 'Gate-Keeping Contextualism'. *Episteme* 8 (1): 83–98.

Henderson, David, and Terence Horgan. 2015. 'What's the Point?'. In *Epistemic Evaluation: Purposeful Epistemology*. Edited by John Greco and David Henderson, 87–114. Oxford: Oxford University Press.

Heraclitus. 1981. *The Art and Thought of Heraclitus*. Edited by Charles H. Kahn. Cambridge: Cambridge University Press.

Hobbes, Thomas. 1839. *The English Works of Thomas Hobbes of Malmesbury*. Edited by Sir Thomas Molesworth. London: John Bohn.

Hobbes, Thomas. 2006. *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiasticall and Civil*. Edited by G. A. J. Rogers and Karl Schuhmann. London: Continuum.

Höffding, Harald. 1887. *Psychologie in Umrissen auf Grundlage der Erfahrung*. Translated by F. Bendixen. Leipzig: Fues.

Holland, Tom. 2019. *Dominion: The Making of the Western Mind*. London: Little, Brown.

Honderich, Ted. 2003. 'Yours Sincerely'. *The Guardian*, 1 March.

Hont, István. 2015. 'Commercial Sociability: The Adam Smith Problem'. In *Politics in Commercial Society: Jean-Jacques Rousseau and Adam Smith*. Edited by Béla Kapossy and Michael Sonenscher, 25–48. Cambridge, MA: Harvard University Press.

Hookway, Christopher. 2010. 'Some Varieties of Epistemic Injustice: Reflections on Fricker'. *Episteme* 7 (2): 151–63.

Horkheimer, Max. 2004. *The Eclipse of Reason*. London: Continuum.

Hoy, David Couzens. 1994. 'Nietzsche, Hume, and the Genealogical Method'. In *Nietzsche, Genealogy, Morality: Essays on Nietzsche's On the Genealogy of Morals*. Edited by Richard Schacht, 249–67. Berkeley, CA: University of California Press.

Hoy, David Couzens. 2009. *The Time of Our Lives: A Critical History of Temporality*. Cambridge, MA: MIT Press.

Huang, Jing. 2019. 'Did Nietzsche Want His Notes Burned? Some Reflections on the *Nachlass* Problem'. *British Journal for the History of Philosophy* 27 (6): 1194–214.

Huddleston, Andrew. 2018. 'Against "Egypticism": Nietzsche on Understanding and "Defining" Concepts'. In *The Nietzschean Mind*. Edited by Paul Katsafanas, 381–94. London: Routledge.

Hufendiek, Rebekka. 2019. 'Das Hypothesenwesen: Die genealogische Methode bei Nietzsche und Rée'. *Deutsche Zeitschrift für Philosophie* 67 (3): 440–50.

Hufendiek, Rebekka. 2020. 'From Natural Hierarchy Signals to Social Norm-Enforcers: What Good Are Functional Explanations of Shame and Pride?'. In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*. Edited by Rebekka Hufendiek, Daniel James, and Raphael Van Riel, 93–121. London: Routledge.

Hufendiek, Rebekka. Forthcoming. 'Beyond Anti-Essentialism: Fine-Tuning Ideology Critique of Appeals to Biology'. *Journal of Social Philosophy*.

Hume, David. 1978. *Essays, Moral, Political, and Literary*. Edited by Eugene F. Miller. Indianapolis, IN: Liberty Fund.

Humeres, Mónica. 2018. 'Crítica a las genealogías reales e imaginarias'. *Cinta de moebio* 63: 314–22.

Hyde, Harford Montgomery. 1973. *The Trials of Oscar Wilde*. New York: Dover.

Iselin, Isaak. 1764. *Über die Geschichte der Menschheit*. Frankfurt am Main/Leipzig: Harscher.

James, William. 1978. *Pragmatism and the Meaning of Truth*. Edited by A. J. Ayer. Cambridge, MA/London: Harvard University Press.

Janaway, Christopher. 2007. *Beyond Selflessness: Reading Nietzsche's Genealogy*. Oxford: Oxford University Press.

Jenkins, Mark P. 2006. *Bernard Williams*. Chesham: Acumen.

Jenkins, Scott. 2016. 'Truthfulness as Nietzsche's Highest Virtue'. *Journal of Value Inquiry* 50 (1): 1–19.

Joas, Hans. 2011. *Die Sakralität der Person: Eine neue Genealogie der Menschenrechte*. Frankfurt am Main: Suhrkamp.

Johnson, Allen W., and Timothy Earle. 1987. *The Evolution of Political Societies: From Foraging Group to Agrarian State*. Stanford, CA: Stanford University Press.

Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge, MA/London: MIT Press.

Kagan, Shelly. 1998. 'Rethinking Intrinsic Value'. *The Journal of Ethics* 2 (4): 277–97.

Kail, Peter J. E. 2007a. 'Leibniz's Dog and Humean Reason'. *Rivista di Storia della Filosofia (1984–)* 62 (3): 65–80.

Kail, Peter J. E. 2007b. *Projection and Realism in Hume's Philosophy*. Oxford: Oxford University Press.

Kail, Peter J. E. 2009. 'Nietzsche and Hume: Naturalism and Explanation'. *The Journal of Nietzsche Studies* 37 (1): 5–22.

Kail, Peter J. E. 2016. 'Hume and Nietzsche'. In *The Oxford Handbook of Hume*. Edited by Paul Russell, 755–76. Oxford: Oxford University Press.

Kant, Immanuel. 1996. *The Metaphysics of Morals*. Cambridge: Cambridge University Press.

Kaplan, Abraham. 2002. *The Conduct of Inquiry: Methodology for Behavioural Sciences*. New Brunswick, NJ: Transaction.

Kappel, Klemens. 2010. 'On Saying That Someone Knows: Themes from Craig'. In *Social Epistemology*. Edited by Adrian Haddock, Alan Millar, and Duncan Pritchard, 69–88. Oxford: Oxford University Press.

Karimzadi, Shahzavar. 2015. *Money and Its Origins*. London: Routledge.

Kavka, Gregory S. 1986. *Hobbesian Moral and Political Theory*. Princeton, NJ: Princeton University Press.

Kelp, Christoph. 2011. 'What's the Point of "Knowledge" Anyway?'. *Episteme* 8 (1): 53–66.

Kersting, Wolfgang. 2017. '"Wohlgeordnete Freiheit". Der Aufbau der Rechtsphilosophie'. In *Vom Naturzustand zur kosmopolitischen Gesellschaft: Souveränität und Staat bei Kant*. Edited by Massimo Mori, 9–38. Wiesbaden: Springer.

Kidd, Ian James, and Havi Carel. 2017. 'Epistemic Injustice and Illness'. *Journal of Applied Philosophy* 34 (2): 172–90.

Kidd, Ian James, José Medina, and Gaile Pohlhaus Jr, eds. 2017. *The Routledge Handbook of Epistemic Injustice*. New York: Routledge.

Kim, Chin-Tai. 1990. 'A Critique of Genealogies'. *Metaphilosophy* 21 (4): 391–404.

Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research*. Cambridge: Cambridge University Press.

Kincaid, Harold. 2020. 'The Empirical Issues in Functional Explanations in the Social Sciences'. In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*. Edited by Rebekka Hufendiek, Daniel James, and Raphael Van Riel, 18–27. London: Routledge.

Kitcher, Philip. 2011. *The Ethical Project*. Cambridge, MA: Harvard University Press.

Klement, Kevin C. 2002. 'When Is Genetic Reasoning Not Fallacious?'. *Argumentation* 16 (4): 383–400.

Kolodny, Niko. 2005. 'Why Be Rational?'. *Mind* 114 (455): 509–63.

Koopman, Colin. 2009. 'Two Uses of Genealogy: Michel Foucault and Bernard Williams'. In *Foucault's Legacy*. Edited by C. G. Prado, 90–108. New York: Continuum.

Koopman, Colin. 2011. 'Genealogical Pragmatism: How History Matters for Foucault and Dewey'. *Journal of the Philosophy of History* 5 (3): 533–61.

Koopman, Colin. 2013. *Genealogy as Critique: Foucault and the Problems of Modernity*. Bloomington, IN: Indiana University Press.

Kornblith, Hilary. 2002. *Knowledge and Its Place in Nature*. New York: Oxford University Press.

Kornblith, Hilary. 2011. 'Why Should We Care about the Concept of Knowledge?'. *Episteme* 8 (1): 38–52.

Korsgaard, Christine. 1996. 'Two Distinctions in Goodness'. In *Creating the Kingdom of Ends*, 249–74. Cambridge: Cambridge University Press.

Krishna, Nakul. 2014. 'The Morality of Common Sense: Problems from Sidgwick'. DPhil Thesis, University of Oxford.

Krishnan, Nikhil. Forthcoming. *A Terribly Serious Adventure: Philosophy at Oxford 1900–60*. New York: Random House.

Krishnan, Nikhil, and Matthieu Queloz. Manuscript. 'The Shaken Realist: Bernard Williams, the War, and Philosophy as Cultural Critique'.

Kusch, Martin. 1995. *Psychologism: A Case Study in the Sociology of Philosophical Knowledge*. London: Routledge.

Kusch, Martin. 2004. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press.

Kusch, Martin. 2009a. 'Review of Epistemic Injustice: Power and the Ethics of Knowing, by Miranda Fricker'. *Mind* 118 (469): 170–4.

Kusch, Martin. 2009b. 'Testimony and the Value of Knowledge'. In *Epistemic Value*. Edited by Adrian Haddock, Alan Millar, and Duncan Pritchard, 60–94. Oxford: Oxford University Press.

Kusch, Martin. 2011. 'Knowledge and Certainties in the Epistemic State of Nature'. *Episteme* 8 (1): 6–23.

Kusch, Martin. 2013. 'Naturalized Epistemology and the Genealogy of Knowledge'. In *Contemporary Perspectives on Early Modern Philosophy: Nature and Norms in Thought*. Edited by Martin Lenz and Anik Waldow, 87–100. Dordrecht: Springer.

Kusch, Martin, and Robin McKenna. 2018. 'The Genealogy of Relativism and Absolutism'. In *Metaepistemology: Realism and Anti-Realism*. Edited by Christos Kyriacou and Robin McKenna, 217–39. New York: Palgrave Macmillan.

Kusch, Martin, and Robin McKenna. 2020. 'The Genealogical Method in Epistemology'. *Synthese* 197 (3): 1057–1076.

Kvanvig, Jonathan L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

Lackey, Jennifer. 2007. 'Why We Don't Deserve Credit for Everything That We Know'. *Synthese* 158 (3): 345–61.

Lane, Melissa. 1999. 'States of Nature, Epistemic and Political'. *Proceedings of the Aristotelian Society* 99 (2): 211–24.

Lange, Friedrich Albert. 1866. *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*. Iserlohn: Baedeker.

Lawlor, Krista. 2013. *Assurance: An Austinian View of Knowledge and Knowledge Claims*. Oxford: Oxford University Press.

Lecky, William Edward Hartpole. 1869. *History of European Morals from Augustus to Charlemagne*. New York: D. Appleton.

Leiter, Brian. 1998. 'The Paradox of Fatalism and Self-Creation in Nietzsche'. In *Willing and Nothingness: Schopenhauer as Nietzsche's Educator*. Edited by Christopher Janaway, 217–57. Oxford: Clarendon Press.

Leiter, Brian. 2015. *Nietzsche on Morality*. 2nd ed. London: Routledge.

Leiter, Brian. Manuscript. 'Williams's Debt to Nietzsche: Real or Illusory?'.

Lifschitz, Avi. 2012. *Language and Enlightenment: The Berlin Debates of the Eighteenth Century*. Oxford: Oxford University Press.

Lillehammer, Hallvard. 2017. 'Smile When You're Winning: How to Become a Cambridge Pragmatist'. In *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Edited by Cheryl Misak and Huw Price, 45–60. Oxford: Oxford University Press.

Locke, John. 2003. *Two Treatises of Government and A Letter Concerning Toleration*. Edited by Ian Shapiro. New Haven, CT: Yale University Press.

Lorenzini, Daniele. Forthcoming. 'On Possibilising Genealogy'. *Inquiry*. 10.1080/0020174X.2020.1712227.

Lottenbach, Hans. 1996. 'Monkish Virtues, Artificial Lives: On Hume's Genealogy of Morals'. *Canadian Journal of Philosophy* 26 (3): 367–88.

Lotter, Maria-Sibylla. 2012. *Scham, Schuld, Verantwortung: Über die kulturellen Grundlagen der Moral*. Frankfurt am Main: Suhrkamp.

Mabille, Louise. 2009. *Nietzsche and the Anglo-Saxon Tradition*. London: Continuum.

Machery, Edouard. 2017. *Philosophy Within Its Proper Bounds*. Oxford: Oxford University Press.

MacIntyre, Alasdair. 1998. *A Short History of Ethics: A History of Moral Philosophy from the Homeric Age to the Twentieth Century*. London: Routledge.

MacIntyre, Alasdair C. 1988. *Whose Justice? Which Rationality?* Notre Dame, IN: University of Notre Dame Press.

MacKinnon, Catharine. 1989. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press.

MacKinnon, Catharine. 1993. *Only Words*. Cambridge, MA: Harvard University Press.

Mandeville, Bernard. 1988. *The Fable of the Bees: or, Private Vices, Publick Benefits*. Edited by F. B. Kaye. Indianapolis, IN: Liberty Fund.

Mannheim, Karl. 1936. *Ideology and Utopia: An Introduction to the Sociology of Knowledge*. London: Routledge.

Margalit, Avishai. 2013. 'Foreword'. In *The Power of Ideas*. 2nd ed. Edited by Henry Hardy, xiii–xxiv. Princeton, NJ: Princeton University Press.

Margolis, Eric, and Stephen Laurence. 1999. *Concepts: Core Readings*. Cambridge, MA: MIT Press.

Margolis, Eric, and Stephen Laurence. 2015. *The Conceptual Mind: New Directions in the Study of Concepts*. Cambridge, MA: MIT Press.

Margolis, Eric, and Stephen Laurence. 2019. 'Concepts'. In *The Stanford Encyclopedia of Philosophy* Edited by Edward N. Zalta. https://plato.stanford.edu/entries/concepts.

Marušić, Jennifer Smalligan. 2017. 'Dugald Stewart on Conjectural History and Human Nature'. *Journal of Scottish Philosophy* 15 (3): 261–74.

McGinn, Colin. 2003. 'Isn't It the Truth?'. *The New York Review of Books* L (6): 70–3.

McGrath, Matthew. 2015. 'Two Purposes of Knowledge-Attribution and the Contextualism Debate'. In *Epistemic Evaluation: Purposeful Epistemology*. Edited by John Greco and David Henderson, 138–57. Oxford: Oxford University Press.

McKenna, Robin. 2014. 'Shifting Targets and Disagreements'. *Australasian Journal of Philosophy* 92 (4): 725–42.

McKenna, Robin. 2015. 'Epistemic Contextualism Defended'. *Synthese* 192 (2): 363–83.

McMahan, Jeff. 2013. 'Bernard Williams: A Reminiscence'. In *The Moral Philosophy of Bernard Williams*. Edited by Alexandra Perry and Chris Herrera, 18–25. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

McMullin, Ernan. 1985. 'Galilean Idealization'. *Studies in History and Philosophy of Science Part A* 16 (3): 247–73.

McShea, Daniel W. 2012. 'Upper-Directed Systems: A New Approach to Teleology in Biology'. *Biology and Philosophy* 27 (5): 663–84.

Medina, José. 2011. 'The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary'. *Social Epistemology* 25 (1): 15–35.

Medina, José. 2013. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*. Oxford: Oxford University Press.

Menand, Louis. 2001. *The Metaphysical Club*. New York: Farrar, Straus and Giroux.

Menge, Torsten. 2017. 'The Uncanny Effect of Telling Genealogies'. *Southwest Philosophy Review* 33 (1): 63–73.

Menger, Carl. 1892. 'On the Origin of Money'. *The Economic Journal* 2 (6): 239–55.

Migotti, Mark. 2016. 'History, Genealogy, Nietzsche: Comments on Jesse Prinz, Genealogies of Morals—Nietzsche's Method Compared'. *Journal of Nietzsche Studies* 47 (2): 212–27.

Millgram, Elijah. 2009. 'D'où venons-nous . . . Que sommes nous . . . Où allons-nous?'. In *Reading Bernard Williams*. Edited by Daniel Callcut, 141–65. London: Routledge.

Millikan, Ruth Garrett. 1989. 'In Defense of Proper Functions'. *Philosophy of Science* 56 (2): 288–302.

Millikan, Ruth Garrett. 2000. *On Clear and Confused Ideas: An Essay about Substance Concepts*. Cambridge: Cambridge University Press.

Millikan, Ruth Garrett. 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press.

Minghella, Anthony. 2000. *The Talented Mr Ripley—Based on Patricia Highsmith's Novel*. London: Methuen.

Misak, Cheryl. 2013. *The American Pragmatists*. The Oxford History of Philosophy. Oxford: Oxford University Press.

Misak, Cheryl. 2015. 'Pragmatism and the Function of Truth'. In *Meaning Without Representation: Essays on Truth, Expression, Normativity, and Naturalism*. Edited by Steven Gross, Nicholas Tebben, and Michael Williams, 262–78. Oxford: Oxford University Press.

Misak, Cheryl. 2016. *Cambridge Pragmatism: From Peirce and James to Ramsey and Wittgenstein*. Oxford: Oxford University Press.

Misak, Cheryl, and Huw Price, eds. 2017. *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Oxford: Oxford University Press.

Möllers, Christoph. 2013. *The Three Branches: A Comparative Model of Separation of Powers*. Oxford: Oxford University Press.

Moore, Adrian W. 1997. *Points of View*. Oxford: Oxford University Press.

Moore, Adrian W. 2006. 'Maxims and Thick Ethical Concepts'. *Ratio* 19: 129–47.

Moore, George Edward. 1993. *Principia Ethica*. Edited by Thomas Baldwin. Cambridge: Cambridge University Press.

Moran, Richard. 2016. 'Williams, History, and "the Impurity of Philosophy"'. *European Journal of Philosophy* 24 (2): 315–30.

Moran, Richard. 2018. *The Exchange of Words: Speech, Testimony, and Intersubjectivity*. New York: Oxford University Press.

Mordacci, Roberto. 2016. 'From Analysis to Genealogy: Bernard Williams and the End of the Analytic–Continental Dichotomy'. *Philosophical Inquiries* 4 (1): 71–84.

Mori, Massimo. 2017. 'Innere Staatssouveränität: zwischen souveräner Unzwingbarkeit und Volkssouveränität'. In *Vom Naturzustand zur kosmopolitischen Gesellschaft: Souveränität und Staat bei Kant*. Edited by Massimo Mori, 101–22. Wiesbaden: Springer.

Moyn, Samuel. 2010. *The Last Utopia*. Cambridge, MA: Harvard University Press.

Mühlebach, Deborah. 2016. 'Reflective Equilibrium as an Ameliorative Framework for Feminist Epistemology'. *Hypatia* 31 (4): 874–89.

Nado, Jennifer. 2019. 'Conceptual Engineering, Truth, and Efficacy'. *Synthese*. 10.1007/s11229-019-02096-x.

Nagel, Ernest. 1977. 'Teleology Revisited: Goal Directed Processes in Biology and Functional Explanation in Biology'. *The Journal of Philosophy* 74 (5): 261–301.

Nagel, Rosemarie, and Fang Fang Tang. 1998. 'Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning'. *Journal of Mathematical Psychology* 42 (2): 356–84.

Nagel, Thomas. 1997. *The Last Word (Philosophical Essays)*. Oxford: Oxford University Press.

Nagel, Thomas. 2009a. 'Williams: Philosophy and Humanity'. In *Secular Philosophy and the Religious Temperament: Essays 2002–2008*, 139–46. Oxford: Oxford University Press.

Nagel, Thomas. 2009b. 'Williams: The Value of Truth'. In *Secular Philosophy and the Religious Temperament: Essays 2002–2008*, 131–8. Oxford: Oxford University Press.

Neander, Karen. 2017. *A Mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, MA: MIT Press.

Nehamas, Alexander. 1985. *Nietzsche: Life as Literature*. Cambridge, MA: Harvard University Press.

Nehamas, Alexander. 2012. 'Nietzsche: Writings from the Early Notebooks'. In *Introductions to Nietzsche*. Edited by Robert B. Pippin, 17–43. New York: Cambridge University Press.

Neuhouser, Frederick. 2014. *Rousseau's Critique of Inequality: Reconstructing the Second Discourse*. Cambridge: Cambridge University Press.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Oxford: Blackwell.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

Oakley, Francis. 1999. *Politics and Eternity: Studies in the History of Medieval and Early-Modern Political Thought*. Köln: Brill.

Ogilvie, Sheilagh. 2019. *The European Guilds: An Economic Analysis*. Princeton, NJ: Princeton University Press.

Origgi, Gloria. 2012. 'Epistemic Injustice and Epistemic Trust'. *Social Epistemology* 26 (2): 221–35.

Owen, David. 2001. 'Wittgenstein and Genealogy'. *Nordic Journal of Philosophy* 2 (2): 5–29.

Owen, David. 2003. 'Nietzsche, Re-Evaluation and the Turn to Genealogy'. *European Journal of Philosophy* 11 (3): 249–72.

Owen, David. 2007. *Nietzsche's Genealogy of Morality*. Stocksfield: Acumen.

Owen, David. 2010. 'Genealogy'. In *Encyclopedia of Political Theory*. Edited by Mark Bevir, 549–51. Thousand Oaks, CA: Sage Publications.

Paden, Roger. 2003. 'Review of Bernard Williams, Truth and Truthfulness: An Essay in Genealogy'. *The Journal of Value Inquiry* 37: 565–9.

Palmeri, Frank. 2016. *State of Nature, Stages of Society: Enlightenment Conjectural History and Modern Social Discourse*. New York: Columbia University Press.

Papineau, David. 2019. 'The Disvalue of Knowledge'. *Synthese*. 10.1007/s11229-019-02405-4.

Pashman, Jon. 1970. 'Is the Genetic Fallacy a Fallacy?'. *Southern Journal of Philosophy* 8 (1): 57–62.

Paul, Laurie Ann. 2012. 'Metaphysics as Modeling: The Handmaiden's Tale'. *Philosophical Studies* 160 (1): 1–29.

Peacocke, Christopher. 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.

Peirce, Charles Sanders. 1931. *Collected Papers of Charles Sanders Peirce*. Edited by C. Hartshorne, P. Weiss, and A. Burks. Cambridge, MA: Belknap Press.

Pettit, Philip. 1996. 'Functional Explanation and Virtual Selection'. *British Journal for the Philosophy of Science* 47 (2): 291–302.

Pettit, Philip. 2000. 'Rational Choice, Functional Selection and Empty Black Boxes'. *Journal of Economic Methodology* 7 (1): 33–57.

Pettit, Philip. 2008. *Made with Words: Hobbes on Language, Mind, and Politics*. Princeton, NJ: Princeton University Press.

Pettit, Philip. 2018. *The Birth of Ethics: Reconstructing the Role and Nature of Morality*. Edited by Kinch Hoekstra. Oxford: Oxford University Press.

Pettit, Philip. 2019. 'Social Norms and the Internal Point of View: An Elaboration of Hart's Genealogy of Law'. *Oxford Journal of Legal Studies* 39 (2): 229–58.

Piccinini, Gualtiero. 2015. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press.

Plunkett, David. 2016. 'Conceptual History, Conceptual Ethics, and the Aims of Inquiry: A Framework for Thinking about the Relevance of the History/Genealogy of Concepts to Normative Inquiry'. *Ergo* 3 (2): 27–64.

Pohlhaus Jr, Gaile. 2014. 'Discerning the Primary Epistemic Harm in Cases of Testimonial Injustice'. *Social Epistemology* 28 (2): 99–114.

Prescott-Couch, Alexander. 2014. 'Williams and Nietzsche on the Significance of History for Moral Philosophy'. *Journal of Nietzsche Studies* 45 (2): 147–68.

Prescott-Couch, Alexander. 2015a. 'Genealogy and the Structure of Interpretation'. *Journal of Nietzsche Studies* 46 (2): 239–47.

Prescott-Couch, Alexander. 2015b. 'Nietzsche, Genealogy, and Historical Individuals'. *Journal of Nietzsche Studies* 46 (1): 99–109.

Prescott-Couch, Alexander. 2017. 'Explanation and Manipulation'. *Noûs* 51 (3): 484–520.

Prescott-Couch, Alexander. Manuscript-a. 'Nietzsche and the Significance of Genealogy'.

Prescott-Couch, Alexander. Manuscript-b. 'Rational Reconstruction: Understanding Others through Misrepresentation'.

Price, Huw. 2008. 'Brandom and Hume on the Genealogy of Modals'. *Philosophical Topics* 36 (2): 87–97.

Price, Huw. 2011. *Naturalism Without Mirrors*. Oxford: Oxford University Press.

Price, Huw. 2013. 'Naturalism without Representationalism'. In *Expressivism, Pragmatism and Representationalism*. Edited by Huw Price, 3–21. Cambridge: Cambridge University Press.

Price, Huw. 2017. 'Ramsey's Ubiquitous Pragmatism'. In *The Practical Turn: Pragmatism in Britain in the Long Twentieth Century*. Edited by Cheryl Misak and Huw Price, 149–62. Oxford: Oxford University Press.

Prichard, Harold Arthur. 1950. *Knowledge and Perception*. Oxford: Clarendon Press.

Prinz, Jesse J. 2002. *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.

Prinz, Jesse J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Pritchard, Duncan. 2012. 'The Genealogy of the Concept of Knowledge and Anti-Luck Virtue Epistemology'. In *Conceptions of Knowledge*. Edited by Stefan Tolksdorf, 159–78. Berlin: De Gruyter.

Queloz, Matthieu. 2016. 'Wittgenstein on the Chain of Reasons'. *Wittgenstein-Studien* 7 (1): 105–30.

Queloz, Matthieu. 2017. 'Does Philosophy Have a Vindicatory History? Bernard Williams on the History of Philosophy'. *Studia Philosophica* 76: 137–52.

Queloz, Matthieu. 2019. 'The Points of Concepts: Their Types, Tensions, and Connections'. *Canadian Journal of Philosophy* 49 (8): 1122–45.

Queloz, Matthieu. 2020. 'Revealing Social Functions through Pragmatic Genealogies'. In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*. Edited by Rebekka Hufendiek, Daniel James, and Raphael Van Riel, 200–18. London: Routledge.

Queloz, Matthieu. Forthcoming-a. 'Choosing Values? Williams contra Nietzsche'. *The Philosophical Quarterly*. 10.1093/pq/pqaa026.

Queloz, Matthieu. Forthcoming-b. 'On Ordered Pluralism'. *Australasian Philosophical Review*.

Queloz, Matthieu. Forthcoming-c. 'The Self-Effacing Functionality of Blame'. *Philosophical Studies*. 10.1007/s11098-020-01479-y.

Queloz, Matthieu, and Damian Cueni. 2019. 'Nietzsche as a Critic of Genealogical Debunking: Making Room for Naturalism Without Subversion'. *The Monist* 102 (3): 277–97.

Queloz, Matthieu, and Damian Cueni. Forthcoming. 'Left Wittgensteinianism'. *European Journal of Philosophy*. 10.1111/ejop.12603.

Ramsey, Frank Plumpton. 1990. 'Truth and Probability'. In *F. P. Ramsey: Philosophical Papers*. Edited by D. H. Mellor, 52–94. Cambridge: Cambridge University Press.

Rasmussen, Dennis C. 2017. *The Infidel and the Professor: David Hume, Adam Smith, and the Friendship That Shaped Modern Thought*. Princeton, NJ: Princeton University Press.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls, John. 1985. 'Justice as Fairness: Political Not Metaphysical'. *Philosophy and Public Affairs* 14 (3): 223–51.

Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.

Reader, Soran, ed. 2005. *The Philosophy of Need*. Cambridge: Cambridge University Press.

Reader, Soran. 2007. *Needs and Moral Necessity*. London: Routledge.

Reginster, Bernard. 2006. *The Affirmation of Life: Nietzsche on Overcoming Nihilism*. Cambridge, MA: Harvard University Press.

Reginster, Bernard. 2013. 'Honesty and Curiosity in Nietzsche's Free Spirits'. *Journal of the History of Philosophy* 51 (3): 441–63.

Reynolds, Steven L. 2017. *Knowledge as Acceptable Testimony*. Cambridge: Cambridge University Press.

Richard, Mark. 2019. *Meanings as Species*. Oxford: Oxford University Press.

Richardson, John. 2004. *Nietzsche's New Darwinism*. New York: Oxford University Press.

Ritter, Joachim, Karlfried Gründer, and Gottfried Gabriel. 1971–2007. *Historisches Wörterbuch der Philosophie*. Basel: Schwabe.

Rorty, Richard. 1989. *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.

Rorty, Richard. 1998a. 'Is Truth a Goal of Inquiry? Donald Davidson versus Crispin Wright'. In *Truth and Progress: Philosophical Papers, Volume 3*, 19–42. Cambridge: Cambridge University Press.

Rorty, Richard. 1998b. *Truth and Progress: Philosophical Papers, Volume 3*. Cambridge: Cambridge University Press.

Rorty, Richard. 2002. 'To the Sunlit Uplands'. *London Review of Books* 24 (21): 13–15.

Rorty, Richard, and Pascal Engel. 2007. *What's the Use of Truth?* Translated by William McCuaig. New York: Columbia University Press.

Rousseau, Jean-Jacques. 1977. 'Discourse on the Origin and Foundations of Inequality among Men'. In *The Discourses and Other Early Political Writings*. Translated by Victor Gourevitch, 111–222. Cambridge: Cambridge University Press.

Roux, Wilhelm. 1881. *Der Kampf der Theile im Organismus: Ein Beitrag zur Vervollständigung der mechanischen Zweckmässigkeitslehre*. Leipzig: Engelmann.

Rowett, Catherine. Manuscript. 'Bernard Williams on Truth and Plato's Republic on Justice: What Are Genealogical Arguments Good For?'.

Russell, Paul. 2008. *The Riddle of Hume's Treatise: Skepticism, Naturalism, and Irreligion*. Oxford: Oxford University Press.

Russell, Paul. 2019. 'Hume's Optimism and Williams's Pessimism: From "Science of Man" to Genealogical Critique'. In *Ethics Beyond the Limits: New Essays on Bernard Williams' Ethics and the Limits of Philosophy*. Edited by Sophie Grace Chappell and Marcel van Ackeren, 37–52. London: Routledge.

Rysiew, Patrick. 2012. 'Epistemic Scorekeeping'. In *Knowledge Ascriptions*. Edited by Jessica Brown and Mikkel Gerken, 270–94. Oxford: Oxford University Press.

Saar, Martin. 2002. 'Genealogy and Subjectivity'. *European Journal of Philosophy* 10 (2): 231–41.

Saar, Martin. 2007. *Genealogie als Kritik: Geschichte und Theorie des Subjekts nach Nietzsche und Foucault*. Frankfurt am Main/New York: Campus.

Saar, Martin. 2008. 'Understanding Genealogy: History, Power, and the Self'. *Journal of the Philosophy of History* 2: 295–314.

Sagar, Paul. 2014. 'Minding the Gap: Bernard Williams and David Hume on Living an Ethical Life'. *Journal of Moral Philosophy* 11 (5): 615–38.

Sagar, Paul. 2017. 'Beyond Sympathy: Smith's Rejection of Hume's Moral Theory'. *British Journal for the History of Philosophy* 25 (4): 681–705.

Sagar, Paul. 2018. *The Opinion of Mankind: Sociability and the Theory of the State from Hobbes to Smith*. Princeton, NJ: Princeton University Press.

Sainsbury, R. M., and Michael Tye. 2012. *Seven Puzzles of Thought: An Originalist Theory of Concepts*. Oxford: Oxford University Press.

Salmon, Wesley C. 1973. *Logic*. Englewood Cliffs, NJ: Prentice-Hall.

Sarasin, Philipp. 2008. *Darwin und Foucault: Genealogie und Geschichte im Zeitalter der Biologie*. Frankfurt am Main: Suhrkamp.

Sauer, Hanno. 2018. *Debunking Arguments in Ethics*. Cambridge: Cambridge University Press.

Sawyer, Sarah. 2020. 'Talk and Thought'. In *Conceptual Engineering and Conceptual Ethics*. Edited by Alexis Burgess, Herman Cappelen, and David Plunkett, 379–95. Oxford: Oxford University Press.

Scharp, Kevin. 2013. *Replacing Truth*. Oxford: Oxford University Press.

Scheibenberger, Sarah. 2016. *Kommentar zu Nietzsches Ueber Wahrheit und Lüge im aussermoralischen Sinne*. Berlin: De Gruyter.

Schmid, Jelscha. Forthcoming. '"Es ist so, weil ich es so mache." Fichtes Methode der Konstruktion'. *Fichte-Studien* 48: 389–412.

Schmitt, Carl. 1996. *The Concept of the Political*. Edited and translated by George Schwab. Chicago, IL: University of Chicago Press.

Schroeter, Laura, and François Schroeter. 2015. 'Rationalizing Self-Interpretation'. In *The Palgrave Handbook of Philosophical Methods*. Edited by Chris Daly, 419–47. New York: Palgrave.

Schulte, Joachim. 2017. 'Surveyability'. In *A Companion to Wittgenstein*. Edited by Hans-Johann Glock and John Hyman, 278–90. Chichester: Wiley-Blackwell.

Schwegler, Albert. 1848. *Geschichte der Philosophie im Umriß: Ein Leitfaden zur Uebersicht*. Stuttgart: Franckh.

Searle, John. 2010. *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford University Press.

Service, Elman R. 1975. *Origins of the State and Civilization: The Process of Cultural Evolution*. New York: W. W. Norton.

Shapin, Steven. 1994. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago, IL: University of Chicago Press.

Shklar, Judith N. 1989. 'The Liberalism of Fear'. In *Liberalism and the Moral Life*. Edited by Nancy Rosenblum, 21–38. Cambridge, MA: Harvard University Press.

Simion, Mona, and Christoph Kelp. Forthcoming. 'Conceptual Innovation, Function First'. *Noûs*. 10.1111/nous.12302.

Simpson, Thomas W. 2012. 'What Is Trust?'. *Pacific Philosophical Quarterly* 93 (4): 550–69.

Skinner, Quentin. 1966. 'The Limits of Historical Explanations'. *Philosophy* 41 (157): 199–215.

Skinner, Quentin. 1969. 'Meaning and Understanding in the History of Ideas'. *History and Theory* 8 (1): 3–53.

Skinner, Quentin. 2002. *Visions of Politics I: Regarding Method*. Cambridge: Cambridge University Press.

Skinner, Quentin. 2009. 'Genealogy of the Modern State'. *Proceedings of the British Academy* 162: 325–70.

Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.

Slavov, Matias. 2019. 'Hume, the Philosophy of Science, and the Scientific Tradition'. In *The Humean Mind*. Edited by Angela M. Coventry and Alex Sager, 388–402. London: Routledge.

Sluga, Hans. 1998. 'What Has History to Do with Me? Wittgenstein and Analytic Philosophy'. *Inquiry* 41 (1): 99–121.

Small, Robin. 2005. *Nietzsche and Rée: A Star Friendship*. Oxford: Oxford University Press.

Smart, John Jamieson Carswell, and Bernard Williams. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.

Smith, Adam. 1977. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Edwin Cannan. Chicago, IL: University of Chicago Press.

Smith, Adam. 2002. *The Theory of the Moral Sentiments*. Edited by Knud Haakonssen. Cambridge: Cambridge University Press.

Smyth, Nicholas. 2017. 'The Function of Morality'. *Philosophical Studies* 174 (5): 1127–44.

Sommer, Andreas Urs. 2003. 'On the Genealogy of the Genealogical Method: Overbeck, Nietzsche, and the Search for Origins'. *Bulletin of the Institute of Classical Studies, Supplement* 49 (S79): 87–103.

Sommer, Andreas Urs. 2019. *Kommentar zu Nietzsches 'Zur Genealogie der Moral'*. Berlin: De Gruyter.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume 1*. Oxford: Oxford University Press.

Srinivasan, Amia. 2011. 'Armchair v. Laboratory'. *London Review of Books* 33 (18): 17–18.

Srinivasan, Amia. 2015. 'The Archimedean Urge'. *Philosophical Perspectives* 29 (1): 325–62.

Srinivasan, Amia. 2019. 'Genealogy, Epistemology and Worldmaking'. *Proceedings of the Aristotelian Society* CXIX (2): 127–56.

Srinivasan, Amia. Manuscript. *The Contingent World: Genealogy, Epistemology, Politics*.

Stern, Fritz. 1956. *The Varieties of History: From Voltaire to the Present*. New York: Random House.

Stewart, Dugald. 1858. 'Account of the Life and Writings of Adam Smith'. In *The Collected Works of Dugald Stewart*. Edited by William Hamilton. Vol. 10, 5–98. Edinburgh: Thomas Constable and Co.

Stich, Stephen P. 1990. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press.

Strawson, Peter. 1963. 'Carnap's Views on Constructed Systems versus Natural Languages in Analytic Philosophy'. In *The Philosophy of Rudolf Carnap*. Edited by P. A. Schilpp, 503–18. La Salle, IL: Open Court.

Strawson, Peter Frederick. 1959. *Individuals*. London: Routledge.

Strawson, Peter Frederick. 1961. 'Social Morality and Individual Ideal'. *Philosophy* 36 (136): 1–17.

Strawson, Peter Frederick. 1974. 'Freedom and Resentment'. In *Freedom and Resentment and Other Essays*, 1–28. London: Routledge.

Strevens, Michael. 2008. *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

Strevens, Michael. 2013. *Tychomancy*. Cambridge, MA: Harvard University Press.

Stroud, Barry. 2006. 'The Constraints of Hume's Naturalism'. *Synthese* 152 (3): 339–51.

Teichmüller, Gustav. 1864. *Studien zur Geschichte der Begriffe*. Berlin: Weidmann.

Thatcher, David S. 1989. 'Zur Genealogie der Moral: Some Textual Annotations'. *Nietzsche-Studien* 18: 587–99.

Thomas, Alan. 2008. 'The Genealogy of Epistemic Virtue Concepts'. *Philosophical Papers* 37 (3): 345–69.

Thomas, Alan. 2012. 'Nietzsche and Moral Fictionalism'. In *Nietzsche, Naturalism, and Normativity*. Edited by Christopher Janaway and Simon Robertson, 133–59. Oxford: Oxford University Press.

Thomas, Alan. 2018. 'Should Epistemic Injustices Be Redressed by the "Corrective Virtues"?'. In *Justice: The Virtues*. Edited by Mark LeBar, 209–36. Oxford: Oxford University Press.

Thomasson, Amie L. 2020. 'A Pragmatic Method for Normative Conceptual Work'. In *Conceptual Engineering and Conceptual Ethics*. Edited by Alexis Burgess, Herman Cappelen, and David Plunkett, 435–58. Oxford: Oxford University Press.

Thornhill, Randy, and Nancy Thornhill. 1983. 'Human Rape: An Evolutionary Analysis'. *Ethology and Sociobiology* 4 (3): 137–73.

Thornhill, Randy, and Nancy Thornhill. 1992. 'The Evolutionary Psychology of Men's Coercive Sexuality'. *Behavioral and Brain Sciences* 15: 363–421.

Thucydides. 2013. *Thucydides: The War of the Peloponnesians and the Athenians*. Edited by Jeremy Mynott. Cambridge Texts in the History of Political Thought. Cambridge: Cambridge University Press.

Tieffenbach, Emma. 2011. 'Invisible-Hand Explanations'. Doctoral Thesis, Department of Philosophy, University of Geneva.

Tieffenbach, Emma. 2013. 'Invisible-Hand Explanations: From Blindness to Lack of We-ness'. *Social Science Information* 52 (3): 450–70.

Tolstoy, Leo. 2014. *Anna Karenina*. Translated by Marian Schwartz. Edited by Gary Saul Morson. New Haven, CT: Yale University Press.

Tomasello, Michael. 2016. *A Natural History of Human Morality*. Cambridge, MA: Harvard University Press.

Trestman, Michael A. 2012. 'Implicit and Explicit Goal-Directedness'. *Erkenntnis* 77 (2): 207–36.

Tuck, Richard. 1979. *Natural Rights Theories: Their Origin and Development*. Cambridge: Cambridge University Press.

Tuck, Richard. 1987. 'The "Modern" Theory of Natural Law'. In *The Languages of Political Theory in Early-Modern Europe*. Edited by Anthony Pagden, 99–122. Cambridge: Cambridge University Press.

Tuck, Richard. 1993. *Philosophy and Government 1572–1651*. Cambridge: Cambridge University Press.

Turri, John. 2016. 'The Point of Assertion Is to Transmit Knowledge'. *Analysis* 76 (2): 130–6.

Ueberweg, Friedrich. 1866. *Grundriss der Geschichte der Philosophie von Thales bis auf die Gegenwart, III: Die Neuzeit*. Berlin: Mittler & Sohn.

Ullmann-Margalit, Edna. 1997. 'The Invisible Hand and the Cunning of Reason'. *Social Research* 64 (2): 181–98.

Van Fraassen, Bas C. 2002. *The Empirical Stance*. New Haven, CT: Yale University Press.

Van Riel, Raphael. 2020. 'In Search of the Missing Mechanism: Functional Explanation in Social Science'. In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*. Edited by Rebekka Hufendiek, Daniel James, and Raphael Van Riel, 70–92. London: Routledge.

Voltaire. 1877–85. *Œuvres complètes*. Edited by Louis Moland. Paris: Garnier.

Wanderer, Jeremy. 2012. 'Addressing Testimonial Injustice: Being Ignored and Being Rejected'. *Philosophical Quarterly* 62 (246): 148–69.

Weber, Max. 1922. 'Wissenschaft als Beruf'. In *Gesammelte Aufsätze zur Wissenschaftslehre*, 524–55. Tübingen: Mohr Siebeck.

Weinberg, Jonathan M. 2006. 'What's Epistemology For? The Case for Neopragmatism in Normative Metaepistemology'. In *Epistemology Futures*. Edited by Stephen Hetherington, 26–47. Oxford: Oxford University Press.

Weisberg, Michael. 2007. 'Three Kinds of Idealization'. *The Journal of Philosophy* 104 (12): 639–59.

Weisberg, Michael. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Wierzbicka, Anna. 2018. 'I Know: A Human Universal'. In *Epistemology for the Rest of the World*. Edited by Masaharu Mizumoto, Stephen Stich, and Eric McCready, 215–50. New York: Oxford University Press.

Wiggins, David. 1998. 'What Is the Force of the Claim That One Needs Something?'. In *Necessary Goods: Our Responsibilities to Meet Others' Needs*. Edited by Gillian Brock, 33–55. Oxford: Rowman and Littlefield.

Wiggins, David. 2002a. 'Claims of Need'. In *Needs, Values, Truth*. 3rd ed., 1–57. Oxford: Clarendon Press.

Wiggins, David. 2002b. 'An Indefinibilist cum Normative View of Truth and the Marks of Truth'. In *What Is Truth?* Edited by R. Shantz. Vol. 1, 316–32. Berlin: De Gruyter.

Wiggins, David. 2005. 'An Idea We Cannot Do Without'. In *The Philosophy of Need*. Edited by Soran Reader, 25–50. Cambridge: Cambridge University Press.

Wiggins, David. 2006. *Ethics: Twelve Lectures on the Philosophy of Morality*. Cambridge, MA: Harvard University Press.

Wiggins, David, and Sira Dermen. 1987. 'Needs, Need, Needing'. *Journal of Medical Ethics* 13: 62–8.

Wild, Markus. 2008. 'Hume über das Wahrnehmungsobjekt'. In *Sehen und Begreifen. Wahrnehmungstheorien der Neuzeit*. Edited by Dominik Perler and Markus Wild, 287–317. Berlin: De Gruyter.

Wild, Markus. 2011. 'Hume on Force and Vivacity. A Teleological-Historical Interpretation'. *Logical Analysis and History of Philosophy* 14: 71–88.

Wild, Markus. 2014. 'Wer den Pavian versteht…Eine naturalistische Perspektive auf Wissen bei Mensch und Tier'. *Studia Philosophica* 72: 105–29.

Wild, Markus. Manuscript. *Philosophie und Fiktion: Methodologische Überlegungen zur Rolle von Fiktionen*.

Williams, Bernard. 1973a. 'Deciding to Believe'. In *Problems of the Self*, 136–51. Cambridge: Cambridge University Press.

Williams, Bernard. 1973b. 'Egoism and Altruism'. In *Problems of the Self*, 250–65. Cambridge: Cambridge University Press.

Williams, Bernard. 1973c. 'The Idea of Equality'. In *Problems of the Self*, 230–49. Cambridge: Cambridge University Press.

Williams, Bernard. 1973d. 'Morality and the Emotions'. In *Problems of the Self*, 207–29. Cambridge: Cambridge University Press.

Williams, Bernard. 1981a. 'Conflicts of Values'. In *Moral Luck*, 71–82. Cambridge: Cambridge University Press.

Williams, Bernard. 1981b. 'Internal and External Reasons'. In *Moral Luck*, 101–13. Cambridge: Cambridge University Press.

Williams, Bernard. 1986. 'Reply to Simon Blackburn'. *Philosophical Books* 27 (4): 203–8.

Williams, Bernard. 1988. 'Formal Structures and Social Reality'. In *Trust: Making and Breaking Cooperative Relations*. Edited by Diego Gambetta, 3–13. Oxford: Blackwell.

Williams, Bernard. 1993. *Shame and Necessity*. Berkeley, CA: University of California Press.

Williams, Bernard. 1995a. 'Ethics'. In *Philosophy 1: A Guide through the Subject*. Edited by A. C. Grayling, 545–82. Oxford: Oxford University Press.

Williams, Bernard. 1995b. 'Must a Concern for the Environment Be Centred on Human Beings?'. In *Making Sense of Humanity and Other Philosophical Papers, 1982–1993*, 233–40. Cambridge: Cambridge University Press.

Williams, Bernard. 1995c. 'The Point of View of the Universe: Sidgwick and the Ambitions of Ethics'. In *Making Sense of Humanity and Other Philosophical Papers, 1982–1993*, 153–71. Cambridge: Cambridge University Press.

Williams, Bernard. 1995d. 'Replies'. In *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Edited by J. E. J. Altham and Ross Harrison, 185–224. Cambridge: Cambridge University Press.

Williams, Bernard. 1995e. 'Saint-Just's Illusion'. In *Making Sense of Humanity and Other Philosophical Papers, 1982–1993*, 135–50. Cambridge: Cambridge University Press.

Williams, Bernard. 1995f. 'Truth in Ethics'. *Ratio* 8 (3): 227–38.

Williams, Bernard. 1995g. 'What Has Philosophy to Learn from Tort Law?'. In *The Philosophical Foundations of Tort Law*. Edited by David G. Owen, 487–98. Oxford: Oxford University Press.

Williams, Bernard. 1997a. *Der Wert der Wahrheit*. Wien: Passagen.

Williams, Bernard. 1997b. 'Shame, Guilt, and the Structure of Punishment'. In *Scham und Schuld: Festschrift aus Anlass der Verleihung der Dr. Margrit Egnér-Preise 1997 zum Thema Scham und Schuld*, 21–30. Zurich: Margrit Egnér-Stiftung.

Williams, Bernard. 1999. 'Seminar with Bernard Williams'. *Ethical Perspectives* 6 (3–4): 243–65.

Williams, Bernard. 2000. 'Naturalism and Genealogy'. In *Morality, Reflection, and Ideology*. Edited by Edward Harcourt, 148–61. Oxford: Oxford University Press.

Williams, Bernard. 2001a. 'Liberalism and Loss'. In *The Legacy of Isaiah Berlin*. Edited by Mark Lilla, Ronald Dworkin, and Robert Silvers, 91–103. New York: New York Review of Books.

Williams, Bernard. 2001b. *Morality: An Introduction to Ethics*. Cambridge: Cambridge University Press.

Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005a. 'Conflicts of Liberty and Equality'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 115–27. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005b. *Descartes: The Project of Pure Enquiry*. London: Routledge.

Williams, Bernard. 2005c. 'From Freedom to Liberty: The Construction of a Political Value'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 75–96. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005d. 'Human Rights and Relativism'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 62–74. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005e. 'In the Beginning Was the Deed'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 18–28. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005f. 'The Liberalism of Fear'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 52–61. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005g. 'Modernity and the Substance of Ethical Life'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 40–51. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005h. 'Pluralism, Community and Left Wittgensteinianism'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 29–39. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005i. 'Realism and Moralism in Political Theory'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 1–17. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005j. 'Toleration, a Political or Moral Question?'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 128–38. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2005k. 'Truth, Politics, and Self-Deception'. In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Edited by Geoffrey Hawthorne, 154–64. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006a. 'Descartes and the Historiography of Philosophy'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 257–64. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006b. 'The Elusiveness of Pessimism: Responding to the Ring'. In *On Opera*. Edited by Patricia Williams and Michael Tanner, 62–9. New Haven, CT: Yale University Press.

Williams, Bernard. 2006c. 'Nietzsche's Minimalist Moral Psychology'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 299–310. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006d. 'Philosophy as a Humanistic Discipline'. In *Philosophy as a Humanistic Discipline*. Edited by Adrian W. Moore, 180–99. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006e. 'Plato against the Immoralist'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 97–107. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006f. 'Plato's Construction of Intrinsic Goodness'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 118–37. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006g. 'Political Philosophy and the Analytical Tradition'. In *Philosophy as a Humanistic Discipline*. Edited by Adrian W. Moore, 155–68. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006h. 'The Primacy of Dispositions'. In *Philosophy as a Humanistic Discipline*. Edited by Adrian W. Moore, 67–75. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006i. 'There Are Many Kinds of Eyes'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 325–30. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006j. 'Unbearable Suffering'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 331–40. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006k. 'What Might Philosophy Become?'. In *Philosophy as a Humanistic Discipline*. Edited by Adrian W. Moore, 200–14. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2006l. 'The Women of Trachis: Fictions, Pessimism, Ethics'. In *The Sense of the Past: Essays in the History of Philosophy*. Edited by Myles Burnyeat, 49–59. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2007. 'Truth and Truthfulness'. In *What More Philosophers Think*. Edited by Julian Baggini and Jeremy Stangroom, 130–46. London: Continuum.

Williams, Bernard. 2010. 'Who Needs Ethical Knowledge?'. *Royal Institute of Philosophy Supplement* 35: 213–22.

Williams, Bernard. 2011. *Ethics and the Limits of Philosophy*. Routledge Classics. London: Routledge.

Williams, Bernard. 2012. 'Nietzsche: The Gay Science'. In *Introductions to Nietzsche*. Edited by Robert B. Pippin, 137–51. Cambridge: Cambridge University Press.

Williams, Bernard. 2014a. 'Anarchy, State, and Utopia, by Robert Nozick'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 107–14. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014b. 'The Last Word, by Thomas Nagel'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 371–87. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014c. 'The Need to Be Sceptical'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 311–18. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014d. 'Political Liberalism, by John Rawls'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 326–32. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014e. 'Realism with a Human Face, by Hilary Putnam'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 320–6. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014f. 'Sense and Sensibilia and Philosophical Papers, by J. L. Austin'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 40–5. Princeton, NJ: Princeton University Press.

Williams, Bernard. 2014g. 'Why Philosophy Needs History'. In *Essays and Reviews 1959–2002*. Edited by Michael Woods, 405–12. Princeton, NJ: Princeton University Press.

Williams, Michael. 2013. 'How Pragmatists Can Be Local Expressivists'. In *Expressivism, Pragmatism and Representationalism*. Edited by Huw Price, 128–44. Cambridge: Cambridge University Press.

Williams, Michael. 2016. 'Pragmatism, Sellars, and Truth'. In *Sellars and His Legacy*. Edited by James R. O'Shea, 223–60. Oxford: Oxford University Press.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, Timothy. 2003. 'Understanding and Inference'. *Aristotelian Society Supplementary Volume* 77 (1): 249–93.

Williamson, Timothy. 2005. 'Knowledge, Context, and the Agent's Point of View'. In *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Edited by Gerhard Preyer and Georg Peter, 91–114. Oxford: Oxford University Press.

Williamson, Timothy. 2017. 'Model-Building in Philosophy'. In *Philosophy's Future: The Problem of Philosophical Progress*. Edited by Russell Blackford and Damien Broderick, 159–72. Hoboken, NJ: Wiley-Blackwell.

Williamson, Timothy. 2018a. *Doing Philosophy: From Common Curiosity to Logical Reasoning*. Oxford: Oxford University Press.

Williamson, Timothy. 2018b. 'Model-Building as a Philosophical Method'. *Phenomenology and Mind* 15: 16–22.

Wimsatt, William C. 1972. 'Teleology and the Logical Structure of Function Statements'. *Studies in History and Philosophy of Science Part A* 3 (1): 1–80.

Wittgenstein, Ludwig. 1979. *Notebooks, 1914–1916*. 2nd ed. Translated by G. E. M. Anscombe. Edited by G. H. von Wright and G. E. M. Anscombe. Chicago, IL: University of Chicago Press.

Wittgenstein, Ludwig. 1981. *Zettel*. 2nd ed. Edited by G. E. M. Anscombe and G. H. von Wright. Oxford: Blackwell.

Wittgenstein, Ludwig. 2000. *Wittgenstein's Nachlass: The Bergen Electronic Edition*. Edited by The Wittgenstein Archives at the University of Bergen. Oxford: Oxford University Press.

Wittgenstein, Ludwig. 2009. *Philosophische Untersuchungen = Philosophical Investigations*. Rev. 4th ed. Translated by G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte. Edited by P. M. S. Hacker and J. Schulte. Chichester: Wiley-Blackwell.

Wray, Randall L. 1998. *Understanding Modern Money*. Cheltenham: Edward Elgar.

Wright, Larry. 1973. 'Functions'. *Philosophical Review* 82 (2): 139–68.

Young, Julian. 2010. *Friedrich Nietzsche: A Philosophical Biography*. New York: Cambridge University Press.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Index

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.