

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 anticipatable on the basis of the needs we started out from. Having identified developments 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 appropriate, 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 human purposes 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, *especially 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 other 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 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.