

4

Ideas as Remedies to Inconveniences

David Hume

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

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

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

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

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

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

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

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

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

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

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

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

4.1 Motivating Genealogy: Artificiality and the Circle Argument

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

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

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

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

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

- (P1) The virtue of an action derives entirely from the virtue of the action’s motive.
- (P2) An action must be virtuous before we can have a sense of its morality.
- (C1) The sense of an action’s morality cannot be the *original* motive for the action, since that would be circular (though it does not exclude its later becoming a motive for the action).

⁵ See Cohon (2006) and Kail (2016).

⁶ See Blackburn (1998, 33).

- (C2) Therefore, for any virtuous action, there has to be some virtuous motive—other than the sense of its morality—from which the action derives its virtue (*T*, 3.2.1.4).
- (C3) Therefore, for any virtuous action, there has to be some motive in human nature that originally motivates the action (*T*, 3.2.1.7).

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

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

- (P3) Any motive must be either a natural motive or an artificial motive.
- (P4) In the case of justice, we find no natural motive sufficient to explain the full range of just action.
- (C4) Therefore, there must be an artificial motive for justice that is distinct from the sense of its morality.

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

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

4.2 A Remedy to Conflict Over External Goods

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.3 De-Instrumentalizing Justice

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

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

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

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

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

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

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

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

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

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

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

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

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

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

¹² Smith’s conception of sympathy is broader than Hume’s, however; see Sagar (2017). See Cohon (2010) for a systematic overview of Hume’s moral philosophy.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.4 Promising: Enabling Reciprocal Cooperation Over Time

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

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

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

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

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

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

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

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

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

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

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

4.5 The Functions of the State of Nature

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

²¹ These natural endowments function as unexplained explainers: ‘It is needless to push our researches so far as to ask, why we have humanity or a fellow-feeling with others . . . We must stop somewhere in our examination of causes’ (EPM, 5.2.2n19).

²² Heraclitus (1981, fragment 23). See Williams (1995e, 137–8).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Deliberative Schema A

- (P1) We need to solve the problem of conflict over possessions.
- (P2) Adopting rules that stabilize possessions by turning them into property and cultivating respect for property as a virtue would solve that problem.
- (C1) Therefore, let us adopt these rules and cultivate that virtue.

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

Deliberative Schema B

- (P1) I have reason to abstain from the possessions of others, provided they abstain from mine.
- (P2) Others similarly have reason to abstain from my possessions, provided I abstain from theirs.
- (P3) I know that they know this, and they know that I know that they know this, and so on, which means that it is common knowledge.
- (C1) Therefore, I shall tentatively abstain from the possessions of others.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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