Nietzsche’s Pragmatic Genealogy of Justice

MATTHIEU QUELOZ

Philosophy, University of Basel

This paper analyses the connection between Nietzsche’s early employment of the genealogical method and contemporary neo-pragmatism. The paper has two goals. On the one hand, by viewing Nietzsche’s writings in the light of neo-pragmatist ideas and reconstructing his approach to justice as a pragmatic genealogy, it seeks to bring out an under-appreciated aspect of his genealogical method which illustrates how genealogy can be used to vindicate rather than to subvert and accounts for Nietzsche’s lack of historical references. On the other hand, by highlighting what Nietzsche has to offer neo-pragmatism, it seeks to contribute to neo-pragmatism’s conception of genealogy. The paper argues that Nietzsche and the neo-pragmatists share a naturalistic concern and a pragmatist strategy in responding to it. The paper then shows that Nietzsche avoids a reductive form of functionalism by introducing a temporal axis, but that this axis should be understood as a developmental model rather than as historical time. This explains Nietzsche’s failure to engage with history. The paper concludes that pragmatic genealogy can claim a genuinely Nietzschean pedigree.

Introduction

‘The thinker’, Nietzsche writes, ‘regards everything as having evolved […] he asks: whence does it come? what is its purpose?’ (WS 43). In his account of justice in Human, All Too Human, Nietzsche lets an answer to the second question grow out of an answer to the first: from the naturalist standpoint from which everything is seen as having evolved, a grasp of purpose is sought via genealogy.

This triad of naturalism, pragmatism, and genealogy is echoed in contemporary neo-pragmatism,¹ and it is this parallel which this paper explores and exploits. The paper has two goals. On the one hand, it seeks to bring out an under-appreciated aspect of Nietzsche’s genealogical method by viewing his texts in the light of neo-pragmatist ideas and reconstructing his approach to justice as a pragmatic genealogy. This serves to illustrate how genealogy can be used to vindicate rather than to subvert and accounts for Nietzsche’s neglect of history. On the other hand, the paper seeks

¹ A group whose contemporary representatives include Edward Craig, Hugh Mellor, Cheryl Misak, Robert Brandom, Huw Price, Michael Williams, Simon Blackburn, Robert Kraut and Paul Horwich. For an overview, see Misak (2007) and Misak and Price (forthcoming). Connections between Nietzsche and (classical) pragmatism are regularly drawn by Richard Rorty; see Allen (2010) and Fairfield (2010) for detailed assessments.
to show that neo-pragmatic genealogy has a genuinely Nietzschean pedigree. This serves to indicate what neo-pragmatism might gain from a reading of Nietzsche.

It might seem odd to connect Nietzschean genealogy to neo-pragmatism. After all, a core aim of neo-pragmatism is to accommodate topics of discourse that seem difficult to place within the naturalistic worldview by understanding our sayings in terms of our doings (Price 2011, 231; Brandom 2000, 18, 2008, 2013, 86); where it succeeds, it tends to vindicate our thinking and speaking in certain ways in terms of its practical value for creatures like us. Nietzschean genealogy, by contrast, is still widely seen as being in the business of subverting or problematising its target by highlighting its contingency.2

Yet the genealogy of justice in HA shows that Nietzsche’s method fundamentally grows out of a concern about naturalisability. It answers to a nineteenth-century naturalist’s concern that, without metaphysical postulates, some topics of discourse are hard to make sense of in naturalistic terms. Genealogy remedies this by offering a diachronic translation back into nature of what appeared to be beyond the naturalist’s grasp.

Moreover, Nietzsche seeks to understand the ‘most universal ideas, the last wisps of smoke from the evaporating end of reality’ (TI ‘Reason’, 4) not in terms of the intrinsic nature of their referents, but in terms of the function in human affairs of these ideas themselves. Nietzsche’s focus on functions has received little attention, as interpreters (Foucault 1971; Nehamas 1985; Geuss 1994; Koopman 2013) have tended to understand genealogy as highlighting contingency rather than functional necessity and to assimilate it to history. Nor have functional considerations received much attention in the literature on Nietzsche’s treatment of justice (Knoll 2009; Petersen 2015; Sedgwick 2013; Patton 2008, 2013). An exception is Richardson (2004, 2008), for whom genealogy uncovers the purpose of our practices by viewing them as expressions of our drives and asking what these drives have been selected for.3 Yet by focusing on drives, Richardson neglects the role of concepts in explaining the lure of metaphysics which Nietzsche seeks to dispel. I argue, by contrast, that Nietzsche explains the appeal of metaphysics in terms of the concepts we live by, and these concepts in terms of our needs.

This places Nietzsche in a tradition of pragmatic naturalism culminating in contemporary neo-pragmatists like Huw Price. Price, like Rorty, has claimed Nietzsche

---

2 See Koopman (2013) for an overview.
3 Another is May (1999), who also reads Nietzsche’s genealogies as fictional but valuable ‘ways of getting us to think, even if hypothetically, about the functions of our actual ethical practices and their motivations by relating them to possible earlier or more elementary practices and motivations in a manner which is free of the search for timeless “groundings”’ (52).
as a precursor, though without substantiating the claim (2011, 186, 2013, i, 5).\textsuperscript{4} What such roughly Wittgensteinian forms of pragmatism have to offer Nietzsche is a clear articulation of the idea that metaphysical worries about a given notion can be dissipated by looking at its role in practice and exhibiting its relation to the needs of creatures like us.\textsuperscript{5}

However, this species of functionalist explanation runs up against the fact that what we do when we use a term tends not to reduce to the performance of well-delineated functions. Various contingencies are likely to have left their mark on our conceptual practices besides functional dynamics. If we focus on drives, we run the danger of occluding that fact by only encouraging an instrumental view of conceptual practices as subservient to drives. If we focus on concepts, we have more room for the idea that history imbied them with non-functional aspects, so that their roles fail to line up with their functions. The challenge then becomes to illuminate the role by the function without reducing one to the other.

It is in his appreciation of the importance of taking a diachronic perspective that Nietzsche is of particular interest to contemporary neo-pragmatism.\textsuperscript{6} Nietzsche was among the first to realise that by introducing a temporal axis, we can avoid reducing the concept at issue to a functionalist understanding of it which eclipses its non-functional features. As another contemporary pragmatist put it, ‘[w]ith genealogy, we need neither overstress nor overlook function’ (Craig 2007, 198).

Viewing Nietzschean genealogy in a pragmatist light has another advantage: the interpreters mentioned so far all tend to read Nietzsche as writing some form of history. Yet this raises the problem of accounting for what can only appear, in a trained philologist, as a poor effort at writing history. This problem, real enough in the Genealogy, becomes even more acute in the case of the genealogy of justice: its only historical reference is to a situation in which questions of justice precisely failed to arise. If, by contrast, we read Nietzsche as trying to identify the original function of justice using a kind of model—what I call a type situation—his failure to engage with history becomes comprehensible.

\textsuperscript{4} Another precursor being Hume, whose own genealogy of justice (2000, 3.2.2) bears a striking resemblance to Nietzsche’s. See Hoy (1994).

\textsuperscript{5} See Misak (2016) for an illuminating account of the connections between Wittgenstein and pragmatism. For a pragmatist reconstruction of Wittgenstein’s views on reasons and rationality, see Queloz (2016).

\textsuperscript{6} Unsurprisingly—given the prevalence of Darwinism and historicism in the nineteenth century—an emphasis on temporality is also characteristic of the classical pragmatists and of some neo-pragmatists (see Koopman 2009). However, the neo-pragmatists at work today tend to emphasise functionality at the expense of temporality; and the diachronic element in Nietzsche I am concerned with here is not history, but the time-axis of a developmental model.
In sum, this paper argues for the following claims: (i) Nietzsche’s genealogical method grows out of a concern to naturalise the seemingly metaphysical; (ii) it addresses this concern by making sense of it in terms of its practical value; (iii) characteristic of this strategy is the shift from the phenomenon to our concept of it; (iv) Nietzsche avoids reductive functionalism by introducing a temporal axis; (v) this temporal axis is not that of history, but that of a developmental model. I will spell out each of these claims, illustrate them with a reconstruction of Nietzsche’s genealogy of justice, and end by discussing what such a fictional genealogy achieves.

**Genealogy as Diachronic Translation into Nature**

A central question for Nietzsche, with which he opens not only *HA*, but also *Beyond Good and Evil*, is how something can emerge from its opposite. Nietzsche’s talk of ‘opposites’ is, as he himself suggests (*HA* 1), hyperbolic. The question is not so much whether something can emerge out of its *opposite* in nature, but whether it can emerge out of the *rest* of nature (thus the rational is to be derived not so much from the *irrational* as from the *non-rational*). Nietzsche’s concern can be expressed as follows: Can *X* emerge from non-*X*, in the sense of everything up to, but not including, *X*? Can we find a place for such ethereal phenomena as *rationality*, *sentience*, *logic*, *altruism*, *truth*, *truthfulness* (*HA* 1, *BGE* 2)—and, we might add, *justice*—in the natural world by explaining how they could have emerged out of the rough-and-tumble of a reality originally devoid of these things?

Philosophers, Nietzsche finds, have tended to answer this question in the negative (*TI* ‘Reason’, 5; *HA* 1). Instead of trying to explain how a highly valued *X* could emerge from a lesser non-*X*, ‘metaphysical philosophy has hitherto surmounted this difficulty by denying that the one originates in the other’ (*HA* 1). Attempts to trace back the ‘supposedly miraculous’ to the ‘complex, the multiply caused’ (*HA* 136) are resisted, even more so after Darwin. ‘Formerly’, Nietzsche writes, ‘one has sought the feeling of the grandeur of man by pointing to his divine origin; this has now become a forbidden way, for at its portal stands the ape, together with other gruesome beasts, grinning knowingly as if to say: no further in this direction!’ (*D* 49).

To pursue the inquiry into origins nonetheless is to attempt to understand a phenomenon as part of nature by understanding it in terms of the *rest* of nature. It is to this end that Nietzsche turns to genealogy in *HA*. Genealogy serves the naturalistic aim of *translating humanity back into nature* (*BGE* 239)—or, as Nietzsche later and rather less grandly put it, of sticking human beings back among the animals (*AC*
Genealogy serves ‘to naturalise humanity with a pure, newly discovered, newly redeemed nature’ (GS 109), whose redeemer, of course, is Darwin.7 If Nietzsche urged philosophers to stop viewing the world sub specie aeterni, it was because he wanted them to view it sub specie evolutionis (HA 2; WS 43; eKGB 1885, 38[14]).

Nietzschean genealogy seeks to show that what appears explicable only by incurring further commitments, paradigmatically of an ontological kind, is intelligible in relation to what is already part of one’s understanding of nature. This is an expression of one of Nietzsche’s guiding methodological principles: the ‘law of parsimony’ (eKGB 1872, 23[30]). ‘Method’, he maintains in BGE, ‘must essentially be the economy of principles’ (13). He articulates this principle already in 1872: ‘the hypothesis which deploys the smallest number of presuppositions and means to explain the world takes precedence over all rivals’ (eKGB 1872, 23[30]). Explanations in terms of ‘simpler and better understood forces, especially of the mechanical sort’ (eKGB 1872, 23[30]), should be given precedence over explanations in terms of more complex or less understood forces. As Bernard Williams (1993) has highlighted, this does not carry with it the demand that the terms in which the explanation is given be the same in every case, as they would be if we tried to describe everything in terms of physics. Rather, what is taken as given will change from one case to the next. Taken in this sense, the law of parsimony boils down to the following demand:

The Law of Parsimony: For any \( X \), \( X \) is to be explained as far as possible in terms we use anyway for everything up to, but not including, \( X \).

To give a genealogy of justice, on this account, is to naturalise justice by explaining it not in terms of physics, but in terms of the rest of human psychology. It is to show how it could have arisen, under specific circumstances, out of motives we take to be effective anyway elsewhere (such as the interest in self-preservation), thereby rendering superfluous the appeal to additional entities (such as special faculties of moral intuition, a realm of forms, or divine commands). Genealogy thus enables the understanding of \( X \) in terms of non-\( X \) by delineating how \( X \) could have emerged from non-\( X \). The relation between \( X \) and non-\( X \), puzzling from a synchronic perspective, is rendered intelligible from a diachronic perspective. A genealogy is a form of diachronic translation back into nature.

By effecting a diachronic translation, Nietzsche paves the way for a non-reductionist naturalism which circumvents issues such as whether \( X \) reduces to or

---

7 Nietzsche’s criticisms of Darwin should not mislead us: Darwinism, he wrote, is something ‘I hold to be true’ (eKGB 1872, 19[132]). Nietzsche knew Darwin mostly second-hand, and many of his criticisms have the effect of radicalising his Darwinism (Richardson 2004, 16–17). Emden (2014) supports this.
supervenes on non-X. The claim is only that X could have emerged out of non-X, in answer to needs human beings have anyway.

**Pragmatic Genealogy**

There is a sense in which Nietzsche’s naturalism forms the problem to which his pragmatism forms the answer. Naturalism provides certain constraints on what is to count as a satisfactory explanation—notably the law of parsimony—and pragmatism offers the strategy by which to satisfy these constraints.

‘Pragmatism’ is a term that has gone through so many hands since it was first coined that it is in danger of losing its embossing. Yet in tracing the seemingly ahistorical and unconditioned to the mundane satisfaction of human needs, Nietzsche’s genealogies substitute philosophical anthropology for metaphysics in a way that is characteristic of pragmatism—they show how, as William James put it, the trail of the human serpent is over everything (1978, 37). Huw Price offers a less gnomic statement of this thought:

Pragmatism begins [...] with phenomena concerning the use of certain terms and concepts, rather than with things or properties of a non-linguistic nature. It begins with linguistic behavior, and asks broadly anthropological questions: How are we to understand the roles and functions of the behavior in question, in the lives of the creatures concerned? What is its practical significance? Whence its genealogy? [...] if we can explain how natural creatures in our circumstances naturally come to speak in these ways, there is no further puzzle about the place of the topics concerned, in the kind of world described by science. (2011, 231–32)

As this and related work by neo-pragmatists suggests, a pragmatist approach involves at least two ideas:

**Ascent to the Conceptual**: In addressing philosophically puzzling phenomena, do not start by asking about the nature or properties of X, but start with the concept of X and the terms in which it is expressed.8

**Pragmatic Direction of Explanation**: Explicate these concepts or terms not by asking what their content or meaning is, but in terms of their role or function in practice, of what we are doing when we think and speak in this way. What does the concept or term do for us, and what is the rationale that drove its adoption?9,10

---

8 See Misak 2007; Blackburn 2013, 71; M. Williams 2013, 128.
9 See Misak 2007; Brandom 2000, 12, 18; Price 2011, 29; Blackburn 2013, 71.
10 It is worth noting that the neo-pragmatists, unlike most classical pragmatists, do not take the pragmatic direction of explanation to entail a commitment to empiricism (see Brandom 2000, 23–25; M. Williams 2013). On this conception of pragmatism, Nietzsche is therefore free to offer fictional stories rather than empirical histories of the origins of concepts.
These two ideas are recognisably (if largely implicitly) at work in Nietzsche’s thought. Reading him in the light of neo-pragmatist work can help us make these ideas explicit and lend support to Nietzsche’s approach.

The debt is not entirely one-sided, however. Nietzsche’s work can help us reinforce an aspiration which, though prominent in classical pragmatism (Koopman 2009), is comparably neglected in contemporary neo-pragmatism: the aspiration to add a diachronic dimension to pragmatic explanation. Though I do not have room to argue for it here, some neo-pragmatists tend to run together two questions one can ask of a concept: what it now does, which is a matter of its current role in human affairs and likely covers a wide range of disparate employments; and what its original function is. Nietzsche, especially in his later work, appreciates that the answers to these two question are unlikely to line up: concepts are subject to reinterpretation, they can lose or acquire new functions; crucially, however, reinterpretations often only ‘obscure’ (GM 2.12) but do not erase past functions. As a result, our current concepts form an amalgam of numerous functions that have been layered into them over the course of history.\(^{11}\) Moreover, contingent historical developments are likely to have imbued them with aspects that cannot be captured in functionalist terms (GM 1.2). This raises a problem: on the one hand, the point of a concept is what the naturalist would really like to identify, because this is what explains how creatures in our circumstances naturally came to think in terms of it (not as a causal explanation of why it first arose, but as a functional explanation of why, having arisen, it endured); on the other hand, our current practices are the products of so much history that they are unlikely to exhibit a shape that lends itself to a simple and purely functionalist understanding—they will be in various ways multi- and non-functional.

Nietzsche recognises that the way to deal with this gap between current role and original function is to add a third idea to his pragmatism, namely

**Diachronic Orientation:** The pragmatic account must be given along a diachronic axis, in the form of a *genealogical explanation* of how and why we might have come to think and speak in these terms. The crucial thing for Nietzsche, as for contemporary neo-pragmatists, is not to answer these questions in ways that encourage metaphysics.\(^{12}\)

Introducing a diachronic axis allows Nietzsche to connect our current ways of thinking with their functional origins without reducing them to this original functionality. He emphasises this desideratum in the *Genealogy* (1.2, 2.12–14), but even in *HA*, where he makes sense of justice in terms of its instrumental value, he already leaves room for the thought that its current value goes beyond the instrumental.

\(^{11}\) This point, present only *in nuce* in *GM* 2.12, becomes central in 2.13 and 2.14.

\(^{12}\) See *HA* 10, Price (2007, 95) and Blackburn (2013, 69).
A similarly diachronic pragmatism is on display in Nietzsche’s approach to the seemingly unconditioned values of morality: he asks what the ‘value of those values’ (GM P 6) is, trying to determine their function and decide whether they promote and enhance life by investigating their possible origins (BGE 4); and in his approach to the ‘puzzling’ phenomenon of unconditioned truthfulness, or what he calls ‘the will to truth’, he asks about the ‘value of this will’ (TL 1). Nietzsche’s genealogies naturalise the seemingly unconditioned by presenting our thinking in such terms as functional responses driven by needs. As he puts it: ‘Our concepts are inspired by our need’ (eKGWB 1885, 2[77]).

In revealing concepts we think of as unconditioned by history and functionality to be in fact thus conditioned, Nietzsche’s genealogies do for concepts what Darwin has done for organic traits. ‘If there is something new in Nietzsche’s use of genealogy’, Clark writes, ‘it is the suggestion that concepts are formed in the same way as other living things’ (2015, 31). A useful trait’s emergence will often be accidental; but its stability and spread through a population will be non-accidental: it will endure because it is connected to something useful. Nietzsche accounts for the spread of concepts by the same logic: there is variation of cultural formations in the course of history, and the usefulness of a given formation helps explain its retention and perpetuation.

Nietzsche takes concepts to be inadequate tools for copying the world, because they originate from the obfuscation of differences and involve ‘false’ but life-serving abstractions, but they are necessary for coping in the world: ‘with this invented and rigid world of concepts and numbers, man gains a means of seizing by signs, as it were, huge quantities of facts and inscribing them in his memory’ (eKGWB 1885, 34[131]). Some concepts may be better at fulfilling this function than others. Over time, the concepts that earn their keep will persist or even spread, while those that do not will go out of business. This is why ‘the most useful concepts have remained; however wrong their origin may have been’ (eKGWB 1885, 34[63]).

If we see the concepts we live by in this light, we will be less receptive to the seductions of grammar that encourage metaphysics (BGE P, 20). A note puts it succinctly: ‘Up to now, one generally trusted in one’s concepts as a miraculous dowry from some miracle world: but in the end they were the legacies left us by our most distant, stupidest and yet cleverest forebears;’ ‘concepts and words are our inheritance from days when heads were very dim and modest’ (eKGWB 1885, 34[195]). It is a model of these days when heads were very dim and modest that Nietzsche’s genealogies take as their point of departure.

For a detailed account of Nietzsche’s pragmatic genealogy of truthfulness, see Queloz (forthcoming).
The Logic of Type Situations

If we think that concepts are no miraculous dowry, but the mundane products of pragmatic pressures, then one way to make sense of them is to reconstruct in response to which pressures they emerged. The guiding idea will be to let our understanding of the concept grow out of our understanding of why the concept enjoys such widespread use.

For this type of enterprise, we are less interested in the details of the situation of emergence than in what we might call, in Karl Popper’s phrase, the ‘logic of the situation’ (1957, 149): a reconstruction of the situation in terms of the needs and interests of the agents taking part in it. Such a reconstructed situation will abstract from a host of circumstantial detail to focus on the interplay between the needs and interests involved.

Yet Nietzsche’s point of departure is even more abstract. On the Popperian model, we abstract away from the particulars of a historical situation towards its logic, but we remain concerned with a particular situation and its personalities—what we might call a ‘token situation’ involving ‘token agents’. Nietzsche abstracts from the particular situation altogether. What he ends up with is the logic of type situations involving Weber-like ‘ideal types’ of agents whose needs and interests are articulated in terms of a generic psychology. We might say, therefore, that Nietzsche is concerned with the logic of type situations.

Using such a type situation, we can identify and set out, in rough outline, problems to which human beings at any time and place will need to respond. If we can sketch how practical necessity will drive the emergence of concepts, beliefs and practices that function as solutions to these problems, and if these functional prototypes bear some resemblance to the phenomena we are interested in, we can lay claim to having identified the rationale that drove the emergence of the concepts. This is best understood with the help of a concrete example.

The Genealogy of Justice

In a chapter of HA entitled ‘Of the History of the Moral Sensations’, we find what both the section heading and the context suggest is a genealogy of justice:

Origin of justice.—Justice (fairness) originates between parties of approximately equal power, as Thucydides correctly grasped (in the terrible colloquy between the Athenian

14 Similarly, Janaway argues that Nietzsche’s procedure ‘involves a projected or imagined generic psychology, not properly localized to times, places, or individuals’ (2007, 11).
15 Ullman-Margalit (1977) offers a game-theoretic articulation of this idea.
and Melian ambassadors): where there is no clearly recognizable superiority of force and a contest would result in mutual injury producing no decisive outcome the idea arises of coming to an understanding and negotiating over one another’s demands: the characteristic of exchange is the original characteristic of justice. Each satisfies the other, inasmuch as each acquires what he values more than the other does. One gives to the other what he wants to have, to be henceforth his own, and in return receives what one oneself desires. Justice is thus requital and exchange under the presupposition of an approximately equal power position: revenge therefore belongs originally within the domain of justice, it is an exchange. Gratitude likewise.—Justice goes back naturally to the viewpoint of reasonable self-preservation, thus to the egoism of the reflection: ‘to what end should I injure myself uselessly and perhaps even then not achieve my goal?’—so much for the origin of justice.

Since, in accordance with their intellectual habit, men have forgotten the original purpose of so-called just and fair actions, and especially because children have for millennia been trained to admire and imitate such actions, it has gradually come to appear that a just action is an unegoistic one: but it is on this appearance that the high value accorded it depends; and this high value is, moreover, continually increasing, as all valuations do: for something highly valued is striven for, imitated, multiplied through sacrifice, and grows as the worth of the toil and zeal expended by each individual is added to the worth of the valued thing.—How little moral would the world appear without forgetfulness! A poet could say that God has placed forgetfulness as a doorkeeper on the threshold of the temple of human dignity. (HA 92)

As Nietzsche’s dashes indicate, this section is divided into four parts: (1) the emergence and original function of justice; (2) the original motive to justice; (3) forgetting, imitation, increase in value; (4) the importance of forgetfulness for morality more widely. Let us explore each in turn.

(1) The emergence and original function of justice: the first step in Nietzsche’s approach is so obvious as to be easily missed, but it already does some of the work—it is to ask after the origin of justice, and thereby to historicise a notion which, most evidently in the natural law tradition, presents itself as ahistorical. Against this tradition, Nietzsche maintains that ‘there is no such thing as eternal justice’ (HA 53). Justice has origins, and Nietzsche’s aim is to explain these origins so ‘that it can be perfectly understood without the postulation of metaphysical interference’ (HA 10). This is where he aligns himself with the neo-pragmatist aspiration to explain without encouraging metaphysics. Justice is, as the book’s programmatic title has it, human, all too human. When we ask questions about justice, or about morality generally, ‘we do not touch upon the “nature of the world in itself”; we are in the realm of ideas’ (HA 10). ‘It is’, as he later puts it, we ‘who really and continually make something that is not yet there’ (GS 301). Consequently, the question for Nietzsche is how we came to think in terms of justice, and this question must be ‘relinquished to the physiology and history of the evolution of organisms and concepts’ (HA 10).

This exemplifies the pragmatist ascent to the conceptual: rather than to ask after
the nature of justice, Nietzsche asks how we came to live by the concept of justice. Mankind, Nietzsche writes, ‘set up in language a separate world beside the other world […] man has for long ages believed in the concepts and names of things as in aeternae veritates […] language is, in fact, the first stage of the occupation with science’ (HA 11).

The concept of justice, Nietzsche tells us, originally arises between two parties A and B under the following conditions:

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

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

When these two conditions are met, the most reasonable resolution of the situation for both parties lies in negotiating a settlement through the exchange of desired goods.16,17 This in turn requires the identification of the specific exchange in which ‘[e]ach satisfies the other, inasmuch as each acquires what he values more than the other does.’ Nietzsche’s suggestion is that the concept of justice originated out of the need to negotiate a settlement between equally powerful parties: ‘Der Charakter des Tauschs ist der anfängliche Charakter der Gerechtigkeit’—justice first manifests itself in the just exchange, the exchange that is satisfactory to both parties. Consequently, the original function of justice is to resolve stand-offs between parties of equal power in a manner advantageous to both.

Justice, on this account, is originally justitia commutativa, commutative justice. It is a matter of exchanges being mutually satisfactory to both parties. Lack of satisfaction might then generate a demand for restorative justice towards the injured party and retributive justice towards the injuring party. Hence Nietzsche’s suggestion that ‘revenge therefore belongs originally within the domain of justice, it is an exchange’, and ‘[g]ratitude likewise’: gratitude arises when one is given more than would be just, the desire for revenge when one is given less. This explains why Nietzsche goes on to write that justice is not only ‘exchange’, but also ‘requital’.

When either of the two conditions fails to be met, there can be no question of identifying an exchange acceptable to both parties. If C2 is not met because A

---

16 I use this rationalistic representation of the decision-making process merely as an expository device; it is not meant to describe Nietzsche’s conception of agency.

17 Whether the two parties in fact manage to achieve the most reasonable outcome is then a further question. At least when the two parties in question are nation states, the historical record seems to license scepticism in that regard. See Graham Allison’s work on the ‘Thucydides Trap’.
and $B$ have non-conflicting interests, there can be no question of identifying an exchange acceptable to both parties, since they can satisfy their interests without it. Similarly, if $C1$ is not met, the stronger party will take what it wants and the weaker will have to put up with it. This is the force of Nietzsche’s reference to Thucydides: Thucydides describes how Athens sought to conquer the island of Melos (2013, 5.85–111). The Athenians sent emissaries to the rulers of Melos and offered them an ultimatum: surrender and pay tribute to Athens or be destroyed. The Athenians refused to argue over the justice of the situation, because ‘in the human sphere judgements about justice are relevant only between those with an equal power to enforce it […] the possibilities are defined by what the strong do and the weak accept’ (5.89). This is the Thucydidean insight Nietzsche refers to.

Justice cannot have originated in interactions between parties of unequal power, because, as Richard Crawley rendered the same passage in his 1874 translation, ‘the strong do what they can and the weak suffer what they must’ (1998, 5.89). Hence Nietzsche’s remark in the Genealogy that equilibrium is the presupposition of all contracts ($GM$ 1.4).

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

Justice, then, originates ‘as a settlement between approximately equal powers’ ($GM$ 1.4), as a means to preserve equilibrium rather than to engage in a costly fight for dominance. It emerges as a practical solution to a practical problem. But whose problem and when did it emerge? Nietzsche’s only historical reference is to a case were issues of justice failed to arise. For this reason, it seems clear that Nietzsche is not primarily interested in the specifics of the historical situation in which the concept of justice first arose. Rather, he is concerned with what we called a type situation, characterised by the two conditions we have so far identified.

(2) The original motive to justice: what makes it reasonable for both parties to negotiate is the interest in self-preservation, and this is what Nietzsche brings out in the second part of his genealogy. The original motive to justice in its first configuration is self-preservation rather than a concern for justice itself. This is the
naturalisation of justice, which renders it intelligible by presenting it as growing out of basic human motives.

Moreover, as Nietzsche implies by speaking of ‘reasonable’ self-preservation (*einsichtige Selbsterhaltung*), both parties must be reasonable enough in evaluating their options and their consequences to recognise that an exchange is their best bet. Suicidal agents would presumably forfeit the opportunity to resolve conflicts of interests through exchange, and so would non-suicidal agents too unreasonable to assess the situation correctly. We can therefore enrich our characterisation of the type situation at the origin of justice with two further conditions that specifically concern the type agents:

(C3) *Interest in Self-preservation*: two parties $A$ and $B$ each have a strong interest in self-preservation.

(C4) *Powers of Reasoning*: two parties $A$ and $B$ capable of reasoning their way to the conclusion that given C3, their interests are best served by seeking conflict-resolution through exchange.

That self-preservation forms the original motive to justice is important for two reasons. First, it makes the account *non-circular*, since it explains the emergence of the concept of justice in a way that does not in turn appeal to the concept of justice: one does not originally settle for what come to be thought of as ‘just’ terms of exchange because they are just, but because they are acceptable to both parties and thus instrumental in achieving the desired outcome: the advantageous resolution of conflict. This is connected to Nietzsche’s naturalistic concern to explain the emergence of one thing out of its opposite: he explains the emergence of justice in terms of ‘non-justice’, that is, in terms of considerations that are not themselves matters of justice. Second, it presents justice as emerging out of something we take to be effective anyway elsewhere, thus heeding the demands of parsimony. As Nietzsche puts it in *Daybreak*: ‘The beginnings of justice […] are animal: a consequence of that drive which teaches us to seek food and elude enemies’ (*D* 26).

(3) *Forgetting, Imitation, Increase in Value*: the third part of *HA 92* describes both the veiling of the original function and the development of a substitute conception of justice. The concept of justice spreads and is perpetuated, but not because people reason their way to its instrumental value over and over; rather, people who live by the concept incite their offspring to imitate and admire them. If this goes on long enough, the original function is usually forgotten about and the concept acquires a life of its own: justice becomes an *independently motivating* reason for action. *Ex hypothesi*, utility in a purely personal sense can motivate just acts and has done so. But it also limits the scope of justice by making its exercise conditional on
its having beneficial consequences. In insisting on the importance of the fact that these self-interested origins were forgotten, Nietzsche acknowledges that anything recognisable as our concept of justice is not tied to utility in this way. It is only if the functional origins of justice are veiled and replaced by another motive that the extension of the concept’s scope to parties of unequal power becomes intelligible as a developmental possibility. Nietzsche describes this process most fully in the second volume of *HA*:

> The same actions that within primitive society were first performed with a view to common utility have later been performed by other generations from other motives: out of fear or reverence of those who demanded and recommended them, or out of habit, because one had seen them done all around one from childhood on, or out of benevolence, because their performance generally produced joy and approving faces, or out of vanity, because they were commended. Such actions, whose basic motive, that of utility, has been forgotten are then called moral actions. ([WS 40](#))

The outcome of this development is that people no longer seek just exchanges in confrontations between equals because it is conducive to self-preservation; the justification they give to themselves and to others is that it is just, and the causal explanation for this is that they have been taught to admire such actions as just—a new, non-egoistic motive and reason for action has arisen, and justice is sought for its own sake. Moreover, as Nietzsche argues in the passage on the origin of justice, the value of justice continually increases as the value of the efforts and sacrifices of those who have sought it in the past is added to it. That it was hard-won in the past becomes part of the reason why one should continue to seek it now.

What exactly happens when the original function of justice—that it serves to resolve stand-offs between parties of equal power—is forgotten? People continue to seek mutually satisfactory exchanges, and so—other things being equal—the function of conflict resolution continues to be performed. But it is no longer part of the concept-users’ conception of justice that it performs that function, and consequently they are no longer motivated by that instrumental reason. Though it may remain functional, justice is now thought of in non-functional terms. It comes to be treated as an independently motivating reason for action—that something is just becomes in itself a reason to do it.

(4) *The importance of forgetfulness for morality more widely*: because Nietzsche holds that forgetting is the psychological mechanism responsible for the veiling of the original function of justice, and because he thinks that similar dynamics have given rise to similar moral concepts, he takes forgetfulness to be crucial to morality more widely. If the functionality of such things as justice remained in view, they would continue to be sought for that reason, and non-egoistic or moral
considerations would fail to get a grip: they would either be redundant, because they recommended what prudential considerations recommended anyway; or, where they conflicted with prudential considerations, they would constantly be overridden by them. This is why Nietzsche concludes section 92 with the claim that the world would be a lot less moral without forgetfulness.

There are two ways of reading this claim. On one reading, encouraged by the Foucauldian expectation that genealogies will be disobliging, Nietzsche means that forgetfulness about its functional origins is essential to our allegiance to justice. That self-preservation forms the original motive to justice is important because it differs from our present understanding of justice as something that is precisely not a matter of what Nietzsche describes as ‘egoistic reflection’. It presents justice in merely instrumental terms as a reasonable means towards an egoistic end, and therefore as not quite being what it seems to be.

On the other reading, Nietzsche’s conclusion that the world would be a lot less moral without forgetfulness means that the veiling of functional origins achieved by forgetfulness is a causally necessary step in the development of moral concepts, but without being logically necessary to the finished product—much as scaffolding can be necessary to erecting a house, but where this does precisely not entail that the finished product will fall to pieces once it is removed. The claim that many moral concepts could never have arisen if their origins had not been forgotten then means that this was necessary for them to acquire a value that went beyond the instrumental. This non-instrumental value once acquired, however, the concepts are stabilised by moral motives, independently of whether or not they are assisted by prudential motives.

In fact, once justice has become a moral notion valued for its own sake, this may well raise suspicions of its own: some may feel uneasy about justice, suspecting it to be a mysterious or fetishised notion. Nietzsche is conscious of this, and it is one reason why he thinks we need genealogy: the liberation through which the philosopher ‘emerges from superstitious and religious concepts’ and ‘overcome[s] metaphysics’ is one thing—‘Then, however, a retrograde movement is necessary: he has to grasp the historical justification that resides in such ideas, likewise the psychological; he has to recognize that they have been most responsible for the advancement of mankind and that without such a retrograde movement one deprives oneself of the best that mankind has hitherto produced’ (HA 20). Genealogical inquiry is this retrograde movement. Far from always subverting its target, genealogy can serve to vindicate it by revealing it to be a response to human needs.
The Point of Fictional Genealogy

All this would be poor history if Nietzsche were trying to write history. But there are other ways of approaching questions of origins. Nietzsche’s approach to justice is helpfully compared to that adopted by ethologists like Konrad Lorenz or Richard Dawkins in explaining restraints on aggression: many animals turn out to be much less aggressive than a naive interpretation of life as a struggle for self-preservation might predict. They often exhibit ritualisations—gestures of surrender, for example, are recognised by victors, who then refrain from dealing the killing blow. Lorenz and Dawkins make sense of this by abstracting away from the particulars of given encounters and deriving the rationale driving the emergence of the behaviour from type situations (Dawkins 1976, 67–73). Like Nietzsche’s confrontation between equally powerful parties, the type situation is, as Dawkins admits, ‘naively simple’—it is ‘a “model,” something that does not really happen in nature, but which helps us to understand things that do happen in nature’ (74). With the help of such a model, Lorenz and Dawkins—like Nietzsche—end up presenting quasi-moral behavioural patterns as rational because functional responses to challenges of self-preservation. This suggests a fictionalist reading of Nietzsche’s genealogies as depicting not what actually happened, but a schematic model of what we could imagine to have happened.

Nietzsche’s genealogy of justice is thus emblematic of what Bernard Williams had in mind when he defined genealogy as ‘a fictional story which represents a new reason for action as being developed in a simplified situation as a function of motives, reactions, psychological processes which we have reason to acknowledge already’ (2000, 159). It presents justice as being developed in a situation combining power equality (C1) with conflicting interests (C2) as a function of the motive of self-preservation (C3) and the capacity to reason one’s way to what best serves one’s interests (C4), and subsequently becoming an independently motivating and genuinely new reason for action.

The point of starting with such a type-situational model is to identify generic initial conditions the dynamics of which drive the emergence of justice. This does not yield a categorical claim about how the concept of justice actually came about. It yields a hypothetical claim about how it could have come about: if such-and-such conditions are fulfilled, then we should expect something illuminatingly described as ‘proto-justice’ to arise. The explanatory power of the genealogy will be directly proportional to the level of generality and abstraction achieved by the initial conditions. The less specific the initial conditions are, the greater the probability that the genealogy describes the dynamics through which justice actually came about.
Moreover, since the explanation is given in functional terms, it also yields a claim about justice’s original function. Here also there is obvious value in spelling out in highly abstract terms the initial conditions describing the exigencies to which justice functions as a remedy. The more culturally specific and parochial the circumstances to which a concept is presented as constituting a functional response, the less likely that the genealogy has identified a major driving force behind the concept’s development. Presenting a concept as functional relative to certain circumstances also suggests that it will cease to be functional once these circumstances cease to obtain. Hence, the more parochial the circumstances are, the sooner the concept will cease to be functional, and the less its functionality will have contributed to explaining its prevalence.

What counts as a successful execution of a genealogical explanation is a function of the purpose to which it is put. If, as I have suggested, its purpose is to answer concerns about naturalisability, it will be enough to show that something could have arisen in a way that does not encourage metaphysics. Presenting a prototype of the target concept as a natural because functional response to certain exigencies of life lends plausibility to the story, giving even the most austere naturalist reason to be comfortable with the concept.

Yet it is not only in showing the possible origins of the concept of justice that Nietzsche’s genealogy elucidates the modality of the concept; in revealing the concept of justice to be contingent upon certain facts about concept-users and their environment, it also reveals how local these facts are; and where, as in the case of justice, the facts are extremely general, the genealogy exhibits the concept of justice as counterfactually robust: even if many things had been different, we would likely still have had something performing the role of the concept of justice, because wherever conditions C1–C4 obtain, there are strong pragmatic pressures on it to arise. By highlighting the relative practical necessity of certain concepts, genealogy can lay bare constraints on the space of possible concepts we might live by.

To wrap things up, let me address two worries that this reading is bound to elicit. One is that even if Nietzsche is concerned with the generic, this does not exclude his being concerned with the contingent and parochial; he may be talking in generic terms about Europe only. This rightly highlights that something counts as generic only relative to a reference class; the dynamics highlighted by Nietzsche’s genealogy apply only as far as its four initial conditions. But on the fictionalist reading offered here, he is committed only to saying that where these conditions do obtain, a form of justice will arise—not to saying that they actually obtained. His four conditions at the origin of justice are so minimal, however, that they present
justice as approximating practical necessity and universality.

Another worry is that *HA* is representative only of Nietzsche’s positivistic phase, which he overcomes in the *Genealogy* when he condemns the functionalist approach of the ‘English psychologists’ (*GM* 1.2) and criticises as a ‘psychological absurdity’ the idea that one might forget about functionality despite its acting as a constant reminder of itself (*GM* 1.3). While this points to important developments in Nietzsche’s use of genealogy, I think it understates its continuity (evidenced by the fact that in the *Genealogy*’s preface, he still refers his readers to his genealogy of justice in *HA*). Forgetfulness about functionality indeed gives way to a picture in which functionality never enters consciousness in the first place: in his genealogy of the concepts of good and evil, Nietzsche highlights their functionality in securing the revenge of the slaves without presupposing comprehension on their part—the process is ‘unconscious’ (although on a fictionalist reading of the genealogy of justice, this may be no more than a difference in presentation). Moreover, Nietzsche comes to see the need to insist that functionalist abstraction should not be confused with history—first, because this would amount to a reductive simplification of our concepts, and second, because functionalist abstraction can only take us from generic problems to generic responses to them, but cannot account for the historical variations of these responses. Type situations can yield the conclusion that any society needs *some* set of ethical ideas, but when what is at stake is precisely the development that *differentiates* one particular set of ethical ideas from another (that of nineteenth-century Christians from that of the Greeks, for example), references to more contingent historical transformations become necessary. Nevertheless, the Nietzsche of the *Genealogy* holds on to—indeed, becomes more explicit about—the need to ascend to the conceptual (see *GM* 1.2–4); he still works in terms of generic psychology and type situations (though they become more differentiated); and the first step in his inquiry into the origin of the Slave morality and the ascetic ideal is to naturalise them by revealing them to be instrumental to the expression of the will to power of the weak and disadvantaged. It is only after having unriddled how they can be *locally* life-promoting that he then goes on to evaluate whether they are life-promoting *overall*. His genealogy of morality accounts for a historically acquired dysfunctionality, but it does so ultimately in terms of original functionality. Hence, his later use of genealogy can be seen as refining rather than relinquishing his earlier method.
Conclusion

In this paper, I have argued that Nietzsche’s early genealogy of justice is fruitfully viewed as a pragmatic genealogy in the vein of contemporary neo-pragmatism. My reconstruction shows that it is best seen not as responding to historical concerns about the datable origins of justice, but as responding to concerns about naturalisability. It further shows that in alleviating these concerns, Nietzsche’s genealogy pursues the neo-pragmatist strategy of replacing the question of what justice is by the question of the practical value of thinking in terms of justice; he then goes on to answer that question in terms of a narrative which presents the concept of justice as having naturally emerged as a functional response to a highly general problem. As we have seen, Nietzsche’s genealogical method holds out the promise of a non-reductionist naturalism. His genealogical explanation is, initially, a functionalist story, but one which, thanks to its temporal indexing, does not overemphasise function. It is only the first stage of development which is presented as forming an immediate response to human needs. There is room for further stages, less directly constrained by those needs and more responsive to the contingencies of history, which may carry us far from those functional origins. Genealogy neither overlooks nor overemphasises function.

I want to conclude with an assessment of the significance, for our understanding of Nietzsche and of neo-pragmatism, of reading the account of justice in \textit{HA} as I have done in this paper. Its significance for our understanding of Nietzsche lies in the fact that it offers us an early prototype of his genealogical method which can serve as a guide to his more complex deployment of the method in the \textit{Genealogy} (more complex, in particular, because of its involvement in genuine history). The functionalist developmental model which Nietzsche offers us in \textit{HA} already implies that there is a more complex and genuinely historical story to be told about how less functionally necessitated and more contingent developments turned the prototype notion of justice into the rich, history-laden concept we know today. This kind of development gains in importance in the \textit{Genealogy}, where the desire to make sense, on naturalistic terms, of morality as a generic human phenomenon stands alongside the desire to explain and highlight the differences between its specifically Christian outgrowth and earlier expressions of the phenomenon. In this enterprise, which is essentially that of explaining current dysfunctionality in terms of original functionality, abstract type situations cannot take us all the way. Historical developments need to be taken into account at various points.

This is not to say that Nietzsche’s later version of the genealogical method is necessarily better. For the concerns of naturalisability at the centre of \textit{HA}, no history
NIETZSCHE’S PRAGMATIC GENEALOGY OF JUSTICE

is required. One only needs to show that something ‘can be perfectly understood without the postulation of *metaphysical interference*. This does not exclude its subsequently developing into something altogether different, thus leaving room for the idea that our current concepts may no longer be amenable to a purely functionalist understanding. But the functionalist diachronic model allows us to make sense of them in naturalistic terms, thus rendering metaphysical postulates redundant.  

For neo-pragmatism, the significance of Nietzsche’s genealogy of justice lies in providing us with a common ancestor for seemingly unrelated forms of genealogy and a further historical reference point by which to make sense of the contemporary scene. Two projects in particular emerge as more genuinely Nietzschean than previously supposed. One is Edward Craig’s *Knowledge and the State of Nature* (1990), which offers a diachronic pragmatic account of knowledge that can claim to be a direct heir of Nietzsche’s pioneering exploration of diachronic pragmatism. The other is Bernard Williams’s *Truth and Truthfulness* (2002), which emphasises the need to enrich genealogies given in terms of type situations with history and thereby parallels the development from Nietzsche’s genealogy of justice to his genealogy of morality. Neither Craig’s nor Williams’s genealogical inquiries have the effect of undermining what they are about, but on the account I have offered in this paper, this makes them no less Nietzschean. A Nietzschean genealogy of justice will indeed have a destabilising effect on someone who thought of it in terms of theological or metaphysical revelations. Yet what it destabilises is not one’s commitment to justice, but to metaphysical interpretations of it. And if we had scruples about justice because it appeared mysterious or difficult to integrate into a naturalistic world-view, Nietzschean genealogy will even be vindicatory. Pragmatic genealogy can thus claim a genuinely Nietzschean pedigree.

---

18 Of course, naturalism itself involves assumptions some will consider ‘metaphysical’. But while in one sense, metaphysics is the project of describing reality’s substratum *beyond* the sensible world—the project of Schopenhauer’s *The World as Will and Representation* (1819/1844)—in another sense, it is the project of describing a substratum *internal* to the sensible world—the project of Schopenhauer’s *On the Will in Nature* (1836). While Nietzsche rejects metaphysics in the former sense, he may well embrace metaphysics in the latter sense with his doctrine of the will to power (*BGE* 13, 36, 259; *GS* 349).

19 In *Knowledge and the State of Nature*, Craig mentions neither Nietzsche nor genealogy; but see Craig (2007).
NIETZSCHE'S PRAGMATIC GENEALOGY OF JUSTICE

Bibliography

Works by Nietzsche


WS: The Wanderer and His Shadow. In HA, 301-95.


Other Works


NIETZSCHE'S PRAGMATIC GENEALOGY OF JUSTICE


—. ‘Two Kinds of Genealogy?’. Forthcoming.


