Christine Chwaszcza

Hume and the Social Contract. A Systematic Evaluation

Abstract:

The article systematically explores the compatibility of Hume's political philosophy and contractarianism by reconstructing Hume's criticism of the idea of a social contract. In a nutshell, the dispute concerns the theoretical reconstruction of the establishment and maintenance of normative institutions by individual behavior. At the center of the dispute are questions concerning the philosophical analysis of the normative force of obligatory norms, and the theoretical reconstruction of individual persons' reasons—or motives—for following them. The main part of the article is dedicated to the reconstruction of the philosophical motivations behind the different positions. I will contrast contractarian idealism as a theoretical approach for the study of normative phenomena with Hume's empiricist approach. I will also spell out the metaethical differences between the idea of a hypothetical contract and Hume's rule-consequentialist reconstruction of the source of social and political obligations. Returning to the question of whether one can be both a contractarian and a Humean, the different implications of the two approaches for the theoretical understanding of normative rule-following will be presented. The conclusion is that one cannot be both a contractarian and a Humean. The article ends with a defense of the foregoing analysis against two objections.

Keywords: hypothetical contract; Hume on promising, justice and allegiance; normative rule-following.

1. Introduction

The question raised by Sugden (2009) "Can a Humean be a contractarian?" concerns the theoretical reconstruction of the establishment and maintenance of normative institutions by individual behavior. Hume is mainly known as a critic of contractarianism, but he is a critic who shares certain important assumptions with contractarians—most prominently a conventional account of socio-political institutions and a commitment to metaethical subjectivism of some sort. In order to assess Sugden's question it will, therefore, be necessary to take a closer look at the philosophical core of Hume's criticism of contractarianism, which largely concerns methodology—that is to say, questions about the theoretical understanding of moral phenomena and their philosophical reconstruction or explanation.

At its core Hume's criticism is about the explication of the normative force of obligations and the theoretical reconstruction of (contractarian) reasons and (Humean) motives for following normative rules. Whereas Hume's criticism of 16th and 17th century contractarians must be reconstructed from a 'historical' perspective that presents the philosophical motivations of the different approaches within the context (and limitations) of 17th century philosophy, the theoretical assessment of whether one can be a Humean and a contractarian will be presented in purely systematic terms that can be easily realigned with contemporary debates in moral and political philosophy.

The article is organized as follows: I will start with an outline of contractarianism as a particular methodological approach to political philosophy, which I call *ideal* constructivism (*section 2*). In *section 3*, I will present two major philosophical motives for contractarians' recourse to *idealism* and Hume's objections to them. *Section 4* turns to Hume's rule-consequentialist reconstruction of obligations, which is intended to replace the quasi-procedural 'contract'-model of obligations. *Section 5* presents the crucial elements of Hume's account of normative rule-following, which was strongly criticized first by Adam Smith. *Section 6* comes back to the original question of whether one can be a contractarian and a Humean. Hume's explanatory account of the origins of obligatory norms will be critically assessed and contrasted with the contractarian approach. *Section 7* will address two objections to the foregoing analysis that derive from alternative interpretations of Hume's explanatory account of the origins of morality. Readers already familiar with—or not interested in—16th and 17th century methodological debates might skip *section 3* and move directly to *section 4*.

2. Philosophical Contractarianism as a Methodological Approach

Hume distinguishes between political and philosophical contractarianism.[1] He actually criticizes both, but this article will be restricted to the criticism of philosophical contractarianism.

Philosophical contractarians indeed pursue a genuinely philosophical aim. Characteristically, they defend a particular philosophical approach to social and

---

[1] Hume explicitly drew a distinction between political and philosophical contractarianism in his Essay "Of the Original Contract" (Hume 1985b). The idea of a social contract was widely used in politics by different political groups in England who, however, did not work out a theoretical or philosophical account of the idea. For Hume, political parties were mainly Whigs and Tories, and he explicitly wanted to separate the discussion of the political goals of the Whigs from a discussion of systematic philosophical accounts of contractarianism, partly because he maintained that philosophical analysis should not take political sides, and partly because he thought that contractarianism as such can be used in support for either Whig or Tory positions. Since Hume developed his objections and counter-claims to contractarianism over several decades, it will be necessary to refer also to the *Treatise* (Hume 1978[1740]), the second *Enquiry* (Hume 1975[1777]), and selected additional essays.

political philosophy, which I will call *ideal constructivism*. Paradigmatic representatives of philosophical contractarianism are Hobbes (1994[1651]), Rousseau (1997[1762]) and Kant (1996[1797], part II). The main target of Hume's criticism in "Of the Original Contract", however, is John Locke, whose version of philosophical contractarianism is somewhat special. For Locke departs from contractarian mainstream in some methodological respects, which will be discussed below. Also, his *Second Treatise of Government* (1988[1689]) has always been understood as taking political sides. Nevertheless, it is clearly meant to provide systematic philosophical arguments for the position of the Whigs, and carries the argumentative weight of Locke's defense of "a right to resistance",[2] which is the main target of Hume's political criticisms. As will be seen, however, Hume's criticism holds equally for the other versions of philosophical contractarianism; his arguments in the *Treatise* and the *Enquiry* are entirely general.

Despite the striking differences in the details of analysis, philosophical contractarians share a common methodological approach, *ideal constructivism*. What is *ideal* about ideal constructivism is the belief that certain (though not necessarily all) real-world phenomena can be clearly understood only if one can intellectually grasp an ideal and abstract model or concept of the relevant phenomenon.[3] 'True knowledge', so to speak, requires the intellectual cognition of an abstract ideal.[4] That means that true knowledge cannot be gained by studying phenomena as they present themselves in empirical reality. *Constructivism* maintains that such an ideal and abstract concept is to be analytically construed by means of purely logical or conceptual analysis.[5] The philosophical presuppositions, conditions, and terms of a purely hypothetical model of a social contract, which are the flesh and bones of philosophical contractarianism, in this sense serve an epistemological purpose, and are meant to provide the reader with 'true' understanding of philosophically contested concepts such as 'sovereignty' or 'sovereign', 'political obedience' ('allegiance' in Hume), 'state' or 'statehood', the 'unity of the people'[6] or 'civitas' or 'commonwealth', and 'legal authority'.

---

[2] See Hume 1985b, 469. Locke, as far as I know, is the only philosophical contractarian who argues for a right to resistance.

[3] The term 'idealism' needs some further clarification. A remote paradigm of idealism in the relevant sense is Plato's generation of an 'ideal' *polis* in the *Politeia*, which does not exist in empirical reality; a less remote one is Gallileo's formulation of the law of inertia, which abstracts from external force, i.e. empirical conditions of friction. Hobbes positively argues for his method of a hypothetical generation of a commonwealth; see Hobbes 1966[1656], chap. 1.6. But also Kant explicitly defends the idealist method against objections in his popular essay entitled "On the Common Saying: That May be Correct in Theory, but is of No Use in Practice" (Kant 1996[1793]).

[4] Hume explicitly rejects the very idea of abstract concepts, cf. Hume 1978[1740], bk. I, part I, sect. vii. But there is more to his criticism than just that.

[5] For Hobbes, this means giving a hypothetical account of the generation of a commonwealth by social contract; for Rousseau, it means giving an account of a normatively ideal constitution of a political association; for Kant it consists in presenting a normative criterion for legitimacy of government and positive law.

[6] The term 'people' is here used as a technical political term, which defines the normative status of subjects or citizens.

The methodological approach of ideal constructivism is obviously supported by the acknowledgment that the very phenomena of socio-political philosophy are 'man-made', i.e. cultural or conventional institutions, or—as philosophers in the 17th and 18th century would put it—that they are 'artefacts', or 'artificial' in contrast to 'natural' phenomena. Epistemological, i.e. ideal, constructivism in this sense fits well with the subject matter of cognition, but it must not be mistaken for a quasi-ontological thesis. For the ontological thesis as such can be combined with different epistemological and methodological approaches. Indeed, Hume agrees with contractarians about the conventional nature of socio-political institutions, but differs strongly with respect to their philosophical reconstruction. I will therefore speak of 'ideal' constructivism when I want to emphasize the epistemological commitment of contractarianism, and reserve the term 'socio-political' conventionalism for the idea that socio-political institutions are socio-cultural conventions, or artefacts.

In order to assess Hume's methodological criticism, it is necessary to outline some of the philosophical motivations for ideal constructivism.

3. Philosophical Motives for *Ideal* Constructivism

3.1 The Nature of Socio-political Institutions

The normative and theoretical importance of the contractarian emphasis of the institutional nature of socio-political phenomena is likely to be underestimated, because it seems almost self-evident from a modern point of view.[7] In the 17th and 18th century, though, such an understanding still required strong theoretical defense, not least because of its far-reaching normative implications. Socio-political conventionalism means that political competences are attributed to impersonal institutions rather than to identifiable individual persons. Although it is obvious that socio-political institutions must be represented by natural persons, strictly speaking, it is not those persons who bear the relevant competences, but the 'impersonal' socio-political institution which they represent.[8] Thus, 'sovereignty', i.e. the right of making and executing generally binding laws, is a competence attributed to the socio-political institution of government. This institution must be represented by natural persons—either the King, or Parliament, or a more complex construction that involves both. But those persons do not 'own' sovereign competences because of any personal qualities, but only insofar as they represent the relevant institution. Accordingly, no person 'is', strictly speaking, the sovereign. There is only an institution, or 'office' which is so called. To the extent that the sovereign might still be conceived of as an

---

[7] The recognition that competences and responsibilities of government are tied to a socio-political institution, and not to any natural person, is one of the most striking insights of early-modern contractarians. Today it seems commonplace, and hardly anyone would confuse the office—let's say of the president of the United States—with a natural person—let's say J. F. K.

[8] See Chwaszcza 2012 for a reconstruction of Hobbes's position.

entity or person, it is not a natural, but a legal person, which is distinct from the individual human beings who represent it. Similarly, 'the people' as a political entity is not merely a 'multitude of individuals', but a normative status which is attributed to individuals who stand in certain normative relations to each other and to the institution called 'the sovereign'.

To make a long story short, socio-political institutions are nothing but normative fictions, or as Kant put it, "mere thought entities" (1996[1797], part I, §51). Against the background of this analysis contractarian recourse to idealism seems not at all far-fetched. If socio-political institutions are normative fictions, they do not have empirical 'objects of reference'. Thus, their analysis requires a different intellectual approach, which for contractarians takes the form of an abstract analysis of their function.[9]

Hume, notoriously, raised the objection that contractarian ideas are philosophically far too sophisticated for ordinary people to grasp and, therefore cannot provide the basis of political obedience, or in Hume's terms 'allegiance'. His counter-thesis says that political obedience is grounded in *ad-personam* loyalty to ruling families, i.e. to the persons who represent the institutions, and that it is supported by custom and habit as well as by reflection on the practical usefulness of government.

3.2 The Analysis of Normativity

There is a second consideration in favor of idealism, also criticized by Hume. For contractarians are not only interested in 'construing' socio-political institutions, but also in *critically* assessing their normative foundations.[10]

Contractarians explicitly conceive of the relevant socio-political institutions in normative terms, insofar as they are defined in terms of rights and duties, which are determined by the functional competences and responsibilities that must be attributed to the relevant socio-political institutions. According to contractarians, a critical evaluation of the legitimacy of those rights and duties requires a clarification of the sources of the normative status of the relevant institutions, which partly requires a justification of political obedience, and partly an assessment of the legitimate powers of political government and its limitations. When contractarians defend the idea that the source of both is some form of voluntary agreement, i.e. a social contract, they opt for a quasi-*procedural*

---

[9] Any such study must, of course, also take account of socio-political requirements and obstacles in real-world practice, and therefore cannot be fully detached from an analysis of empirical conditions of social life, unless, like Kant, one prefers to restrict the philosophical analysis exclusively to justificatory aspects. The important point according to contractarians is that the *normative* qualities of socio-political institutions cannot be based on 'experience'. Hobbes's contractarianism in particular is meant to overcome 'Aristotelian' traditions of political philosophy that emphasize the role of experience and prudence ('phronesis' in Aristotle). Hume's empiricist criticism of contractarian idealism starts from a post-Aristotelian tradition of empiricism.

[10] Philosophical contractarians, though, give different weight to analytical and normative aspects. In Hobbes, the analytical aspects still play a major role, whereas Rousseau and Kant rather take the institutional character of socio-political phenomena for granted and emphasize the justificatory aspects of social contract theory.

account for the reconstruction of the relevant obligations[11] that expresses the contractarian idea that the natural, i.e. pre-social normative, status of all individual persons is one of equal freedom. Socio-political obligations and inequalities are justified only to the extent that they can be conceived of as deriving from individual agreement, where this agreement has the form of mutual and self-imposed constraints of individual freedom.

The point of the procedural account is, more precisely, the idea that obligations must be *conceivable* as deriving from certain procedures that express individual agreement, even though no such procedure ever takes place. The conditions and details of the social contract are meant to describe a normative point of view, not an instruction manual for generating valid obligations. With the exception of Locke, philosophical contractarians are not at all reluctant to concede that their commitment to a quasi-procedural account of obligation can only take the form of a 'hypothetical' contract. For they would maintain that no empirically contingent event could ever expose the 'true' conditions and presumptions of a critical normative point of view. That is the very point of ideal constructivism. The recourse to the legal tool of making contracts, in this sense, articulates a conceptual 'model', or mere schema, for making sense of the normative validity of political obligation. It is not meant as outlining empirical conditions of legitimacy that ought to be met in non-ideal real-world scenarios.[12]

Hume, famously, is among the first philosophers who thought that contractarianism does not make any sense at all. In his objection he not only doubts the assumption that the mere *idea* of a hypothetical contract can be a *reason* for the recognition of obligations on the side of either subjects or government.[13] In addition he maintains that any analysis of the normative sources of obligation must reflect the sentiments and motives of actual people for following obligatory rules.[14] This requirement is not just an extra feature of Hume's criticism, but an integral part of his 'empiricist' analysis of normative phenomena. Whereas Hume's moral psychology is meant to be explanatory-descriptive, his overall analysis of the normative force of obligations is an exercise in critical moral theory. His criticism of *hypothetical* proceduralism as well as of normative idealism still articulates standard objections to contractarianism.

The most prominent source of Hume's criticism of contractarianism is his essay "Of the Original Contract" (Hume 1985b), which was primarily addressed to

---

[11] In terms of the Aristotelian contrast between distributive and commutative justice, which still prevails in 16th and 17th century, contractarians defend an account of commutative justice against the view that political and social inequalities derive from personal merit, that is, distributive justice in Aristotle's terms.

[12] Contractarian *idealism* notoriously has little to say about standards of legitimacy for particular non-ideal real-world arrangements, which necessarily must fall below *ideal* standards. Hobbes, Rousseau and Kant—unlike Locke—defended the duty of political obedience and explicitly reject a right to resistance.

[13] Dworkin 1975 has raised a similar objection to Rawls 1971.

[14] Hume's scepticism towards 'reason', thus, concerns not only its efficiency for guiding action. His assumption that ideas are only copies of impressions has the consequence that any explication of 'normativity' must somehow be cast in terms of internal impressions, i.e. *sentiments* or, more broadly, moral *motives*.

Locke. But, as will be argued, Hume's criticism is not restricted to the peculiar 'Lockean' aspects of Locke's contractarianism. Locke departs from the contractarian mainstream in at least two respects. First, he feels the need to provide a positive normative foundation for the contractarian idea that the natural normative status of all persons is one of equal freedom.[15] Although it is not clear whether his theory of property rights fits methodologically well with his contractarian argument, Locke himself is certainly convinced that his account of natural individual rights is fully compatible with a conventionalist account of political government along contractarian lines. Second, Locke is not a wholehearted idealist. In fact, he seems to try to support contractarian proceduralism by making it look more 'empiricist'. He sporadically adds traces of empirical evidence about conditions in a state-of-nature and transforms idealist proceduralism into the idea of an ongoing 'tacit contract'.

The latter modification makes Locke a perfect target of objections that superficially do not seem to apply to ideal contractarianism. Focusing on Locke, Hume delightfully maintains that social contract theory simply fits badly with empirical evidence, period. He strongly rejects the idea of a tacit contract by pointing to the obstacles that people intending to leave their country for good would face (see Hume 1985b, 469ff.). However, Hume's claim that historically political government has been established almost everywhere by conquest, not by contract, and is maintained by force (see Hume 1985b, 470),[16] would be rejected by ideal contractarians, who maintain that their analysis is not about the 'power' but about the 'right' to rule. But then again, as will be explained below, Hume's main criticism concerns the *theoretical explication* of the narrative argument of contractarianism, i.e. the philosophical reconstruction of the normative source of (political) obligation and the analysis of the motives for following normative rules. These criticisms are not restricted to Locke, but touch upon the metaethical core of contactarianism and apply to *ideal* contractarianism even more strongly than to Locke. I will now present these criticisms and Hume's alternative reconstruction of socio-political conventionalism.

4. Hume's Rule-Consequentialist Alternative to Contractarianism

The very core of Hume's empiricist approach to political philosophy concerns the theoretical reconstruction of the sources of normative obligations. He argues that contracts cannot be an original source of obligations because the possibility of making contracts already presupposes the acknowledgment of another obli-

---

[15] Locke's recourse to natural law is probably the reasons why Hume (1985b) characterizes contractarianism as "speculative" and tied to theology.

[16] In normative terms, however, Hume recognizes that power and force are insufficient for the maintenance of political authority; cf. Hume 1978[1740], bk. III, part 2, sect. vii; 1975[1777], sect. IV; fn. 28 below.

gation, namely to keep promises. The latter obligation cannot be conceived of without circularity as being created by contract. Its normative force according to Hume derives from the recognition that having the institution of promising is 'good' because it is 'requisite' for society and social life. He thus replaces the quasi-procedural contractarian conception of the normative force of obligations with a rule-consequentialist account.[17] Hume's discussion is not confined to 'political obedience' ('allegiance'), but also includes duties of justice and 'fidelity', i.e. the duty to keep promises. This sometimes makes the assessment of his arguments a bit difficult because not all arguments are equally convincing for all three types of obligations. But since he repeatedly maintains that the cases of fidelity, allegiance and justice are the same, these difficulties will be ignored.

Hume's criticism of contractarianism is straightforward: the normative force of contracts depends on the acknowledgment of the obligation to keep promises. But this obligation cannot itself be explained in terms of contracting. The details of the argument actually vary somewhat between the *Treatise* and the *Enquiry*.

In the *Treatise* Hume seems to emphasize that contracts cannot *generate* normative obligations out of a normative void, because their normative bindingness itself depends on the acknowledgement that promises *ought* to be kept (see especially Hume 1978[1740], bk. III, part 2, sect. VII; 1985b, 480f.). Obligations generated by contract, in this sense, cannot be reduced to motives of self-interest, or to purely prudential reasons, because making contracts already is or presupposes a normative practice, namely the keeping of promises.[18] Put this way, the criticism denies either that political obedience can be grounded in considerations of mere self-interest, that is, prudence, or that such reasons are strong enough to present a motive for allegiance.

In the *Enquiry*, by contrast, Hume pays more attention to the conclusion of his own argument, already present in the *Treatise*, namely the claim that both, the norms of promising *and* the general norms of justice—which according to contractarians are supposedly established by contracts—derive their normative force from the acknowledgement that they are 'good' insofar as they are necessary for the establishment of any form of society.[19] The social function of allegiance, justice, and fidelity is indeed their most important normative characteristic for Hume, and, more important, he maintains that it is identical for all three types of obligations. It must be noted that according to Hume the substan-

---

[17] See Hume 1978[1740], bk. III, part 2, sect. II, 479; 1975[1777], app. III. From a historical perspective it has to be added that Hume develops rule-consequentialism as an alternative to teleological forms of argument. The special challenge of rule-consequentialism for Hume's account consists in its being a version of *rule*-consequentialism. If one wants to describe his position precisely, one would have to add that the moral force of obligations derives from the sentiment that a spectator would have if he were to reflect upon the good consequences of general rules for social life and the bad consequences of their absence, and perceive the beauty and deformity of the resulting systems.

[18] Contractarians need not disagree *tout court*; cf. Hobbes' discussion of "justice", and the "fool's" claim that breaking contracts might be in accordance with reason if doing so is to one's advantage, in chap. XV of *Leviathan*. Hobbes rejects the fool's reasoning as unsound.

[19] Hume 1978[1740], bk. III, part II, sect. I-VII *passim*; 1975[1777], sect. III; sect. IV.

tive content of obligatory norms plays no special role for their normative force. The substantive content is basically empirically contingent and varies with time and place. The *normative* validity of those norms, whatever their substance, is primarily based on the recognition that 'having rules' is good, and the contrast between life in society and absence of society.

Notwithstanding the different targets of Hume's argument in the *Treatise* and in the second *Enquiry*, the main metaethical point is the same in both presentations. Hume develops a *consequentialist* justification of the normative force of obligations, more precisely a *rule-consequentialist* justification. For Hume readily concedes that the good consequences of fidelity, allegiance and justice derive from their recognition as *general* rules, not from an acknowledgment that each act must yield desirable consequences. He also acknowledges that the relevant norms must be considered strictly binding,[20] which is to say that they allow neither for exemptions in light of particular circumstances or consequences nor for varying degrees of fulfillment.[21] Moreover, with respect to obligations of justice, Hume particularly emphasizes the requirement of impartiality in both their range and their application, and argues that impartiality requires a correction not only of self-interest, but also of the natural feelings of sympathy and benevolence, which according to Hume tend to be partial and to favor persons close to oneself.[22] The normative force of obligations, by contrast, requires a normative perspective that is detached from one's personal perspective, a "general and unalterable standard" of approval and disapproval (Hume 1975[1777], sect. V, part 2, 229). For Hume that standard is constituted by reflection on the common good, i.e. the good for society as a whole.[23]

These characterizations fully coincide with the ordinary understanding of obligations, but they differ considerably from Hume's general conception of virtues—in the terminology of the *Treatise* 'natural virtues'—, which allow for degrees of excellence and deficiency, and the practice of which, according to Hume, requires situative particularist judgment. (Natural) virtues and obligations not only have different normative structures, but they are also grounded in different motives. The motives corresponding to 'natural virtues' are mainly sympathy and benevolence, whereas adherence to obligations is motivated by a sense of duty, which derives from the recognition that obedience to general and strictly binding norms is necessary for the maintenance of society as a whole.

This sense of duty is not an 'original' natural feeling, and certainly not an instinct, but involves some form of reflection in order to overcome the natural tendency toward partiality to oneself and persons close to oneself. Although in the end Hume declares that the sense of duty is still 'natural', insofar as it

---

[20] 'Obligations' are what one 'must' do, what is 'to deon' in Greek. The strictly binding quality of obligations is remarkable for Hume because it does not fit the traditional concept of virtue and vice as that which is praiseworthy and blameworthy to do. Fulfillment of obligations is not praised, but required; and non-fulfillment is not merely blamed, but punished.

[21] Hume 1978[1740], bk. III, part 2; sect. II, 497ff.; sect. VI, 531ff.; 1975[1777], app. III.

[22] Hume 1978[1740], bk. III, part 2, sect. II; 1975[1777], sect. III, part 1; sect. IV; sect. V, part 2.

[23] Hume is not very explicit about what makes a common good 'common'. Most of the time he seems to assume that there is a coincidence of individual interests.

derives from a feeling of humanity, or fellow feeling, which is linked to feelings of sympathy and benevolence, he concedes in the *Enquiry* that those latter feelings need to be 'corrected' in order to become impartial. The motives of adherence to norms of obligations are in this sense 'artificial', which is to say in the present context that they are reflection-dependent. In the *Enquiry*, where the distinction between 'natural' and 'artificial' virtues is dropped, Hume explicitly defends the thesis that "public utility is the *sole* origin of justice, and that reflections on the beneficial consequences of this virtue are the *sole* foundation of its merit" (1975[1777], sect. III, part 1, 183).

Two aspects of this line of argument have to be highlighted. First, there is the acknowledgment that the sense of duty involves reflection, insofar as, and because, impartial obligations require the ability to look at things from a point of view that is detached from one's own interests and one's natural, partial, sentiments.[24] Second, it is to be emphasized that norms of obligation are considered useful to the extent that they are strictly binding, and that the sense of duty gains its 'merit' "in virtue of beneficial consequences" that adherence to strictly binding norms has for society as a whole. Hume is, of course, aware that such an account of the normative force of obligations departs from the common and ordinary understanding of their 'deontological' quality, which holds that obligations are strictly binding because they require one to do what is 'right' for intrinsic reasons, that is to say for morally substantive reasons, irrespective of the consequences that follow.[25]

Hume's rule-consequentialist account is significantly different from the contractarian quasi-procedural reconstruction of the normative force of obligations as an (hypothetical) agreement to self-imposed self-constraint of each person's natural liberty by mutual contracts. The contract metaphor reconstructs the normative force of obligations of political obedience and justice as being constituted by some act of voluntary agreement. For Hume, the idea of equal natural freedom seems to make sense only if understood as 'pre-civilized liberty'. It plays no role for the social morality of civilized nations.[26]

---

[24] Hume (1985b, 479) distinguishes two kinds of moral duties, the first deriving from "natural instinct or immediate propensity", "(t)he *second* kind of moral duties are such as are not supported by any original instinct of nature, but are performed entirely from a sense of obligation, when we consider the necessities of human society, and the impossibility of supporting it, if these duties were neglected. It is thus *justice* or a regard to the properties of others, *fidelity* or the observance of promises, become obligatory, and acquire an authority over mankind. For it is evident, that every man loves himself better than any other person, he is naturally expelled to extend his acquisitions as much as possible; and nothing can restrain him in this propensity, but reflection and experience, by which he learns the pernicious effects of that license, and the total dissolution of society, which must ensue from it. His original inclination, therefore, or instinct, is here checked and restrained by a subsequent judgment or observation. [...] The case is precisely the same with the political or civil duty of *allegiance*, as with the natural duties of justice and fidelity." (Hume 1985b, 480)

[25] When post-Humean philosophers speak of 'deontological' norms, they usually refer to the non-consequentialist characterization of their normative force rather then their quality as strictly binding.

[26] In some passages Hume presents the idea of a social contract as that of an 'original' contract and—maybe with a view to Rousseau's second *Discourse on the Origin of Inequality*—as a historical

Hume's criticism of contractarianism can now be summarized. Contractarians explicate political obligations as mutually agreed self-constraints, but overlook the fact that they always have to presuppose the recognition of fidelity, which itself is an obligation that cannot without circularity be reduced to mutual self-constraint (Hume 1978[1740], bk. III, part 2, sect. VIII; 1985b). If the obligation of fidelity cannot be reconstructed in accordance with the idea of mutual self-constraint, it seems methodologically advisable to search for another source of the normative force of obligations which can be applied to all relevant types of norms, including fidelity. The challenge for Hume is to find a reason, or rather a general human motive, for fidelity, which can also serve as a model for allegiance and justice. Rule-consequentialism is his answer.

17th and 18th century contractarians would probably reply that making a promise *is* imposing an obligation upon oneself, and is considered strictly binding because of what it *is* rather than because of the consequences that the acknowledgment of general and strictly binding norms has for society as a whole. 17th and 18th century contractarians would probably also argue that fidelity is different from justice and political obedience insofar as promises are self-imposed constraints of one's freedom, whereas justice and political obedience include other-imposed obligations, that is to say the acknowledgement that other persons have the right to coerce or force one to behave in certain ways.

From a metaethical point of view, the controversy about the normative sources of general obligations between Hume and the philosophical contractarians is still contested. It boils down to the notorious question whether there is a significant metaethical difference between quasi-procedural and rule-consequentialist accounts of justification. For the normative force of a purely 'hypothetical' contract can consist of nothing other than the reasons why one (hypothetically) ought to make it. That raises the question whether there can be (practical) *reasons* that abstract from consequences, which directly leads into the jungle of debates about the nature of (practical) rationality and reasoning. Personally, I am convinced that (practical) rationality cannot be reduced to consequentialist reasoning, and am accordingly willing to argue that contractarians and Hume pursue *alternative* metaethical approaches. If one opts for the contractarian

---

hypothesis, which he thinks is to some extent plausible if conceived of as a pre-historical event, but totally disconnected from the present: "When we consider how nearly equal all men are in their bodily force, and even in their mental powers and faculties, till cultivated by education; we must necessarily allow that nothing but their own consent could, at first, associate them together, and subject them to any authority. The people, if we trace government to its first origin in the woods and desarts [sic!], are the source of all power and jurisdiction, and voluntary for the sake of peace and order, abandoned their native liberty, and received laws from their equal and companion. [...] If this, then be meant by the *original contract*, it cannot be denied that all government is, at first founded on a contract, and that the most ancient rude combinations of mankind were formed chiefly by that principle. [...] The force, which now prevails, and which is founded on fleets and armies, is plainly political, and derived from authority, the effect of established government. [...] But the contract, on which government is founded, is said to be the *original contract*; and consequently may be supposed too old to fall under the knowledge of the present generation." (Hume 1985b, 467–470)

side, however, one had better replace hypothetical proceduralism by a more advanced account of normative justification.[27]

Hume's rule-consequentialist alternative, however, remains attractive. Still, it has its own well-known methodological weaknesses, which continue to inspire philosophical criticism of rule-consequentialism also along contractarian lines (see Rawls 1958). In order to examine the problems of Hume's alternative account, I will now turn to the descriptive-explanatory aspects of Hume's moral psychology, which are an integral part of his 'empiricist' methodology, and which were already criticized by Adam Smith.

5. Elements of Hume's Moral Psychology

As is well-known, according to Hume all moral phenomena must be explicable as psychological phenomena, which qualify as internal impressions also called moral sentiments. What is of interest in the present context is not so much Hume's general thesis as such, but his particular characterization of the sentiments that underlie the recognition of and motivation for following norms of obligation.

In a nutshell, according to Hume *motives* of political obedience, but also of fidelity and justice, are not grounded in 'reasons' or certain forms of 'reasoning'—as he says, "choice and reflection"—, but in moral feeling, i.e. sentiments, which are strengthened by habit and custom in a manner that will be discussed in *section 6* below. Here, I am concerned with the elements that Hume allows into his explanatory account.

These sentiments, to repeat, are not entirely a-rational. From a motivational perspective the problem of following obligatory rules is, for Hume, that such motives frequently collide with two natural tendencies, with (short-term) self-interest and with partiality for persons close to oneself. These tendencies can only be counteracted by another natural sentiment, which is related to those of sympathy and beneficence, but 'corrected' by judgment or reflection (Hume 1978[1740], bk. III, part 2, sect. II; 1975[1777], sect. V, part 2), that is, by the recognition that those norms are necessary for society. This qualification holds equally for fidelity, justice, and allegiance, although the latter, according to Hume is also supported by feelings of *ad-personam* loyalty towards the ruling families.[28]

---

[27] The most prominent alternative is probably Rawls' "reflective equilibrium".

[28] Obviously, those feelings seem to be intended to be different from mere sympathy, and are qualified as 'opinions' in the essay "Of the first Principles of Government" (Hume 1985a), which pursues a somewhat different line of argument than in Hume 1985b. "Nothing appears more surprizing to those, who consider human affairs with a philosophical eye, than the easiness with which the many are governed by the few; and the implicit submission, with which men resign their own sentiments and passions to those of their rulers. When we enquire by what means this wonder is effected, we shall find, that, as FORCE is always on the side of the governed, the governors have nothing to support them but opinion." (Hume 1985a, 32) [...] "Opinion is of two kinds, to

In sect. V of the *Enquiry*, entitled "Why utility pleases", Hume takes recourse to an *experimentum crucis* in order to show that those reflective sentiments are in another sense quite 'natural', i.e. part of the 'natural' psychological equipment of human beings. 'Natural' here is contrasted with the idea that they are a mere product of culture and social education.[29] The question that the experiment is supposed to decide is presented as follows:

> "The social virtues must, therefore, be allowed to have a natural beauty, which, at first, and antecedent to all precept or education, recommends them to the esteem of uninstructed mankind, and engages their affections. And as public utility of these virtues is the chief circumstance, whence they derive their merit, it follows that the end, which they have a tendency to promote, must be some way agreeable to us, and take hold of some natural affection. It must please either from considerations of self-interest, or from more generous motives and regards." (Hume 1975[1777], sect. V, part 1, 214)

The *experimentum crucis*, then, is the observation that in fact persons forego personal advantage in order to do what promotes the common good, and that such actions are indeed praised and approved. Hume feels confident to infer:

> "Compelled by these instances, we must renounce the theory, which accounts for every moral sentiment by the principle of self-love. We must adopt a more public affection, and allow, that the interests of society are not, even on their own account, entirely indifferent to us. Usefulness is only a tendency to a certain end, and it is a contradiction in terms, that anything pleases as means to an end, where the end itself no wise affects us. If usefulness, therefore, be a source of moral sentiment, and if this usefulness be not always considered with reference to self; it follows, that everything, which contributes to the happiness of society, recommends itself directly to our approbation and good-will. Here is a principle, which counts in great part for the origin of morality: And what need we seek for abstruse and remote systems, where there occurs one so obvious and natural." (Hume 1975[1777], sect. V, part 2, 219)

The problem with the *experimentum crucis* is this. The observation that individuals are willing to promote justice or the common good even at the expense

---

wit, opinion of INTEREST, and opinion of RIGHT. By opinion of interest, I chiefly understand the general sense of advantage, which is reaped from government; together with the persuasion, that the particular government, which is established, is equally advantageous with any other that could easily be settled [...]. Right is of two kinds, right to POWER and right to PROPERTY. What prevalence opinion of the first kind has over mankind, may easily be understood, by observing the attachment which all nations have to their ancient government, and even to those names which have the sanction of antiquity." (Hume 1985a, 33) [...] "Upon these three opinions, therefore, of public *interest*, of *right to power*, and of *right to property*, are all governments founded, and all authority of the few over the many." (Hume 1985a, 34)

[29] Hume characterizes such positions as scepticism, and most likely addresses here as elsewhere "Epicureanism", which he thinks reduces morality entirely to motives of self-interest.

of self-interest does not show "*why* utility pleases", but only that persons can be motivated by reasons other than self-interest. His conclusion, that "everything which contributes to the happiness of society, recommends itself *directly* to our approbation and good-will" (emphasis added), follows only if one presupposes that there are exactly two possible motives for such actions, either self-love or a more public affection. Hume's argument, thus, presents a false contrast by suggesting that there are only two possible explanatory options: the relevant sentiments must be reducible to either self-love or to love of society, or mankind, as a whole, *tertium non datur*. And instead of showing that persons really have the good of society as a whole in mind when they respect norms of obligation, Hume simply stipulates *a priori* that they must do so because nothing that is a means is approved of unless the end that it promotes is approved of too.

Adam Smith strongly opposed both Hume's explanatory account of "why utility pleases" and his rule-consequentialist justification of justice in particular. Smith's criticism in part IV of the *Theory of Moral Sentiments* is one of the two passages in his oeuvre where he refers to an "invisible hand", and maintains that certain social consequences are best understood as *un*intended consequences of individual actions.

> "The same ingenious and agreeable author who first explained why utility pleases, has been so struck with his view of things, as to resolve our whole approbation of virtue into a perception of beauty which results from the appearance of utility. No qualities of the mind, he observes, are approved of as virtuous, but such as are useful or agreeable to either the person himself or to others; and no qualities are disapproved of as vicious but such as have a contrary tendency. And Nature, indeed, seems to have so happily adjusted our sentiments of approbation and disapprobation, to the conveniency both of the individual and of the society, that after the striktest examination it will be found, I believe, that this is universally the case. But still I affirm, that it is not the view of this utility and hurtfulness which is either the first or principal source of our approbation and disapprobation. These sentiments are no doubt enhanced and enlivened by the perception of the beauty and deformity which results from this utility and hurtfulness. But still I say they are originally and essentially different from this perception." (Smith 1976[1790], part IV, chap. 2)

Although the consequences are produced by individual activities, these activities are not motivated by considerations about the 'end', i.e. the social consequences which they bring about, but spring from other motives. In the case of justice, for Smith, this is 'propriety', namely, the motive from which one ought to act if one considered the situation from the perspective of an impartial spectator.[30]

---

[30] Smith's objection, although generally correct, is stronger with respect to some norms than to others. Commonly, the motive for keeping a promise is recognition of the duty that one has taken

Being no less 'empiricist' than Hume, Smith's main argument is based on a couple of observations, which Hume actually accepts. This is, first, the observation that approbation of justice and disapprobation of injustice usually refer to particular acts, not to the usefulness of having general rules. The second observation is that the common understanding of why just acts are approved of and unjust acts disapproved of has something to do with the moral quality of the motivation of the agent for the particular act.[31] An empirical-*explanatory* analysis of *motives* for normative rule-following thus according to Smith, supports the meta-ethical claim that justice is to be identified with what is 'right'. Reflections about the distant end that general adherence to those norms produces, by contrast, play no major role in adherence to justice and the approval of just acts. Smith's observations are indeed widely correct, and philosophers generally acknowledge them to be so, even Hume himself.[32]

From a phenomenal point of view, Smith's criticism reaffirms the ordinary deontological understanding of obligations as not only strictly binding, but also as having *instrinsic* normative force—of some sort.[33]

---

on by giving it (voluntarily), not the recognition that society could not exist without the institution of promising. Similarly, disapproving of an arbitrary infliction of harm on another person—e.g. beating a passer-by with a baseball bat—, is not commonly motivated by the good consequences that a general forbearance of such acts has for society as a whole, but because arbitrarily inflicting harm is considered morally bad. By contrast, respecting the right of primogenitur in hereditary laws is probably supported not for intrinsic reasons, but for traditional or utilitarian reasons. Also, acknowledgment that one has to respect positive law is rarely motivated by a belief that each and every law is just, but often endorsed because of considerations concerning the disadvantage of civil disobedience.

[31] Ironically, Smith accuses Hume of exactly the same mistake of which Hume accuses the contractarians: philosophical abstractness. "This beauty and deformity which characters appear to derive from their usefulness or inconveniency, are apt to strike, in a peculiar manner, those who consider, in the abstract and philosophical light, the actions and conduct of mankind. When a philosopher goes to examine why humanity is approved of, or cruelty condemned, he does not always form to himself, in a very clear and distinct manner, the conception of any particular action either of cruelty or of humanity, but is commonly contented with the vague and indeterminate idea which the general names of those qualities suggest to him. But it is in particular instances only that the propriety or impropriety, the merit or demerit of actions is very obvious and discernible. It is only when particular examples are given that we perceive distinctly either the concord or disagreement between our own affections and those of the agent, or feel a social gratitude arise towards him in the one case, or sympathetic resentment in the other. When we consider virtue and vice in an abstract and general manner, the qualities by which they excite these several sentiments seem in a great measure to disappear, and the sentiments themselves become less obvious and discernible. On the contrary, the happy effects of the one and the fatal consequences of the other seem then to rise up to the view, and as it were to stand out and distinguish themselves from all the other qualities of either." (Smith 1976[1790], part IV, chap. 2)

[32] See Hume 1978[1740], bk. III, part 2, sect. I, 477, 481; in the *Enquiry* Hume offers a hypothetical explanation for what he thinks is only an apparent lack of goal-directed reflection in reaction to unjust acts: "What alone will beget a doubt concerning the theory, on which I insist is the influence of education and acquired habits, by which we are so accustomed to blame injustice, that we are not, in every instance, conscious of any immediate reflection of the pernicious consequences of it." (Hume 1975[1777], sect. III, part 2, 203) For a discussion of an interpretation that positively endorses Hume's explanation see the comment to the second objection below.

[33] It has to be noted that the 'intrinsic' normative force need not necessarily be identified with any particular substantive morality—e.g. an individual rights approach. Smith's 'impartial specta-

More important in the present context, however, is the fact that from a methodological point of view, Smith's objection undermines Hume's *argument* for rule-consequentialism by denying the purely instrumental merit of norms of justice. The problem is not that one cannot figure out a better argument than Hume's, but that Hume's characterization of the relevant sentiments is not convincing. As will become clear in the *section 6*, the claim that reflection is restricted to the comparison of two extreme alternatives—either life in society irrespective of the substantive content of obligatory norms or absence of society—is especially implausible as a general account, if the problem is not about having just any set of norms, but a set of norms that in some sense can be said to qualify as 'right'.

For an 'empiricist' approach, Smith's objection articulates a real theoretical challenge. It questions the internal consistency of Hume's methodological approach. And it also has consequences for an assessment of Hume's explanatory-descriptive account of the origins of morality. The lasting attractiveness of contractarianism is partly motivated by the systematic problem that Smith's objection presents for Hume. To be sure, the objection has been raised not only by Smith, but articulates a standard objection to consequentialism.[34] Since Smith seems to be the one who directs it to Hume's explanatory account directly, however, I will refer to it as 'Smith's objection', and will finally turn to Hume's explanatory account of the origins of morals.

6. Hume's Explanatory-Descriptive Account of the Origins of Morality and Its Compatibility with Contractarianism

Famously, Hume's explanation of the origins of justice compares the emergence, or development, of obligatory norms with the dynamics that synchronize the movements of two men, who pull the oars of a boat.[35] The suggestion seems to be that, like the practice of rowing, normative patterns of interaction evolve through practice and become reinforced by repetition, i.e. custom and habit. The emergence of norms of obligation, so to speak, is a problem of coordination insofar as everyone has an interest in supporting any set of norms that is likely to

---

tor' also represents an intrinsically moral point of view, and a utilitarian like J. St. Mill would maintain that what is especially moral about certain norms is their tendency to promote general happiness.

[34] Among the most prominent presentations are, probably, Ross 1930 and Rawls 1955, both of whom criticize utilitarian consequentialism.

[35] Hume 1978[1740], bk. III, part 2, sect. II, 490; 1975[1777], app. III, 306. As Hume clearly recognizes, the problem is actually more complicated, because the establishment and maintenance of norms of obligations requires coordination not of two men, who are in direct contact with each other, but of multiple persons, who act at a distance from each other; see Hume 1978[1740], bk. III, part 2, sect. VII, 538. This complication might be another reason why Hume thinks that reflection about the common good must play some role.

be supported by others too, because ending up in society is better than failing to establish society, whatever set of norms it is that prevails in the end.

The two men's problem of coordination is intellectually indeed comparatively simple: both men pursue the same end, i.e. moving ahead; and the operative dynamic of selection is a relatively obvious re-enforcement mechanism because success and failure can easily and immediately be identified.

It is obvious that coordination becomes more complicated if there is no common end that persons pursue—that is to say, if the social good promoted by norms of obligation is *un*intended—or if Smith's objection is correct that persons perceive of their actions not as mere means to an end, but in terms of their rightness—that is to say, if ordinary persons think that there is something specifically 'moral' about particular types of actions.

In systematic terms, the crucial point of Smith's objection to Hume is not that ordinary persons' moral beliefs are always right in the sense of being incorrigible, but that normative rule-following cannot be reduced to mere behavioral regularities because it requires assessments by *standards of rightness or correctness*. To say that moral norms are followed because they are considered to be 'right' says that they are acknowledged *because* they are considered to be right and *because* individuals believe (rightly or wrongly) that other persons do agree with their judgment about the 'rightness' of the norms. These two 'becauses' require some sort of 'reasons' or 'standards of moral deliberation' that can be used to decide whether a norm, or its application to a particular case, is justified or not. This is the reason why contractarians take recourse to *ideal constructivism*, and why Smith characterizes the judgments of the impartial spectator as those that a moral person ought to share.

In order to avoid a misunderstanding, it must be emphasized that the 'rightness'-requirement does not imply that persons permanently reflect on or scrutinize their moral views. It says that what makes these views 'moral' is that they can be critically examined when they are challenged, and be either backed up by some sort of reasons or else changed. The 'rightness' requirement, so to speak, maintains that ordinary moral practice is to some extent self-reflective and 'critical', insofar as it implies that moral norms are justifiable by reasons or arguments of some sort.[36] This is a claim about the special role of *morality* within the overall set of norms and social regularities.[37] It derives partly from contrasting the normative force of moral rules with that of legal norms, rules

---

[36] The rightness requirement, however, does not select any particular theory of reflection or moral criticism—contractarian *idealism*, Smith's *impartial spectator* point of view, or utilitarianism, or more modern approaches such as Baier's or Rawls' are all possible candidates. Two approaches though seem excluded. This is, first, traditional moral realism, which is at odds with Humean and contractarian approaches for many other reasons; and, second, such versions of moral sense theory that understand moral perception as unreflected, quasi-sensual, perception.

[37] See among others Baier 1965 for an analysis of the role of morality within the overall set of norms and social regularities, and the claim that its particular role consists in raising the question "Such-and-such is the rule, but is it right?". Systematically related is also what Rawls (1958; 1972) calls the problem of stability, i.e. the requirement that ordinary people would still endorse normative principles if their justification (or origin) were public knowledge. Rawls's discussion, however, is mainly restricted to a criticism of Sidgwickean utilitarianism; Baier's is broader.

of etiquette etc., partly from the observation that normative justification is an integral part of ordinary moral practice. It means that critical morality is an integral part of ordinary moral practice—not just a privilege of philosophers.[38]

Where does Smith's objection leave us with respect to an assessment of Hume's explanatory account of the origins of justice? Hume's assimilations of normative rule-following with behavioral regularities either misses the target because it misconstrues the phenomenon of normative rule-following, or remains theoretically unconvincing because it lacks the psychological resources for modeling the critical and reflective aspects of moral practice.

Philosophical contractarians would probably raise the second objection and insist that moral practice must be informed by normative critical standards of what are 'legitimate' or justified rights and obligations from an *ideal* point of view that abstracts from irrelevant empirical factors.

That brings the debate back to the starting point of the contractarian *idealism* as an attempt to develop a critical normative point of view that rational persons ought to accept. As is well-known, contractarians were not able to establish an uncontested critical point of view—nor was anybody else. But the contractarian idea that some such point is needed has been strongly re-affirmed since, as has been the contractarian idea of equal natural freedom of all persons. Hume's restriction of reflection to a choice between "either society or anarchy, period", however, leaves hardly any space for constructing such a critical normative point of view.

As the contrast stands, Sugden's questions can now be answered: One cannot be a contractarian and a Humean for metaethical reasons, as well as with respect to the cognitive structure of normative rule-following.

7. Reply to Two Objections Against the Analysis

Two possible objections to the foregoing analysis have to be addressed. First is the objection that Hume's criticism of contractarianism is directed primarily against the attempt to found morality purely on *reasons* of self-interest,[39] but it does not deny that self-interested 'behavior' plays an important role in the development of justice. Moral acceptance of modest self-interest is indeed a persistent part of Hume's account because he argues that each person by re-

---

[38] Hume would indeed support such a characterization of morality. Like Smith, he rejects the identification of moral conduct with social conformism or unreflected traditionalism. He is a reformer and naturalist, not an irrationalist. It is the very point of the attempt to prove the influence of utility on moral judgment to provide a criticial standard for the normative assessment of general norms and individual behavior. His attempt to reconcile the project of critical morality with the general outline of the account of moral sentiments and his account of (practical) reason just produces more theoretical difficulties than he seems to have expected, and like many others, he ends up wanting to have his cake and eat it too.

[39] Hume seems to think that Hobbes defends such an argument (Hume 1975[1777], app. II), but it can certainly not be ascribed to Locke and Rousseau, or to Kant, who writes later.

flection can recognize that it is in his interest, i.e. self-interest, to respect norms of obligations because it is requisite for social life. What the rowing example is meant to show, according to the first objection, is that justice can arise by custom and habit from 'reflected' self-interest. But the processes which yield such a result are *not* directly intended as an end. Would Smith's objection still apply to Hume's explanatory account?

The alternative interpretation changes the outline of what would be Hume's argument, for it must now be asked: "Under what conditions, *given Hume's assumptions about people's motivations*, can it be expected that self-interested behavior leads to normative coordination by custom and habit?"

Hume's assumptions, to repeat, are the following: (1) Men are naturally inclined towards being partial to oneself and to others close to oneself, unless those inclinations are corrected by reflection and checked by other moral sentiments. It is clear that Hume agrees with (2): There is no unique set of obligatory norms that enables individuals to enjoy the advantages of society, but a plurality of such sets. It also seems permissible to assume that Hume would agree with (3): Different sets of norms tend to favor different groups in society. If these conditions hold, the coordination problem again becomes more complicated than the rowing example suggests.[40] I assume that under such a description, it would be rather implausible to suppose that coordination emerges purely by habit and custom irrespective of the substantive content of the relevant norms, and that Smith's objection would still apply.

It is, of course, an interesting project to analyze the conditions under which an invisible-hand mechanism would coordinate individual patterns of strategic choice given assumption (1) to (3) above, and also how likely it is that such conditions prevail. For explanatory purposes, however, merely identifying the possibility of such coordination is not enough. One also has to ask how likely it is that such coordination is at work compared with other mechanisms, be they power and force, or moral discourse.

Experience at least seems to tell us that under most conditions it is most likely that persons will quarrel about which set of norms should be established. Unless their quarrels are decided by power and force, they will have to come up with some form of general reason why one set of norms is better justified than

---

[40] In game-theoretic terms, the problem is that multiple non-equivalent equilibria exist; the situation would resemble an n-person chicken game; see Taylor 1987. Little, it seems, can be said about such a problem in the abstract, because the question whether games with multiple equilibria have a unique solution, and under what conditions, seems to depend on the properties of particular games (pay-offs, strategies, combinations of strategies) and the kind of model (standard, iterated, evolutionary, etc.). There is, however, a striking difference between game theorists and Hume. In general, traditional game theorists seem interested in the question whether it is (individually) 'rational' to follow normative rules, which is not Hume's question. Evolutionary game theory, which suspends 'choice' and uses algorithms as strategies, seems in some respects closer to Hume's explanatory interests. I will not comment on the question whether it makes sense to use algorithms as representations of 'motives' for normative rule-following for analytical reasons. For explanatory purposes, it has to be acknowledged that if motives for moral behavior cannot be reduced to mere behavioral regularities, then *a forteriori* they cannot be reduced to algorithms.

some of the alternatives. That does not imply that choice between competing sets of normative rules will, or must, be settled exclusively by the force of better argument. But it is highly unlikely that any normatively acceptable settlement will occur without critical moral discourse about the alternatives.

The conclusion again seems to be that one cannot be a contractarian and a Humean. 'Justice', as Rawls rightly maintains, is about principles for settling conflicts of interest. It is, so to speak, not about how to play the game, but about which rules to accept for playing the game from a critical normative point of view. It is unlikely that any rules from a plurality of options will be accepted without reflection and exchange of justificatory reasons.

The second objection arises from a sceptical background and would question the relevance of Smith's objection to Hume's argument. For a Humean might simply maintain that the understanding of obligations by ordinary people is itself already a product of historical—or, as Mill (1861, chap. 4) would maintain, of psychological—mechanisms of education and acculturation.[41] Real-world persons, as Mackie (1977) claims, are simply in error about the true sources of their normative beliefs. Hume's explanatory account of the origins of morality, therefore, need not presuppose that moral feelings are already at work in the establishment of norms, but only in their maintenance. Once a particular set of norms has emerged for empirically contingent reasons, the substantive content of those norms becomes draped with superstitious beliefs, with the result that people (erroneously) come to think that they are 'right' for intrinsic reasons.[42]

In order to discuss the second objection, one has to ask: "What would such a society look like?" "Under what conditions could persons hold superstitious normative beliefs?". Given the description, it must be assumed that its members would hold rather homogenous moral beliefs, and that moral practice would be 'conservative', i.e. strongly resistant to criticism and (demands of) change. Although it cannot be excluded that such an 'explanation' of the origins of morals might be true for some societies at some points of time, the question arises, which, or how many, societies are like that.[43] The picture certainly does not fit the nature and structure of moral practice in modern societies, which is much

---

[41] Such an interpretation is suggested by some passages in Sugden 1986. A similar, but still different line of interpretation is given by Haakonssen (1981) who argues that Hume in the *Treatise* tries to explain the historical development of justice as an unintended effect of 'efficient causes', by which he means particular actions by individual persons, whereas the normative force (in Haakonssens terminology "the moral quality") of justice is explicated in terms of goal-directed reasoning. Haakonssen, however, argues that Hume's explanatory account fails, because no mere aggregation or accumulation of particular actions by different persons can constitute a 'general' rule in Hume's own terms.

[42] Hume 1978[1740], bk. III, part 2, sect. V; 1975[1777], sect. III, part 2, 199, where Hume compares rules of justice and superstitious beliefs as follows: "But there is a material difference between superstition and justice, that the former is frivolous, useless, and burdensome: the latter is absolutely requisite to the well-being of mankind and society. When we abstract from this circumstance (for it is too apparent to be overlooked) it must be confessed, that all regards to right and property, seem entirely without foundation, as much as the grossest and most vulgar superstition."

[43] For reasons explained above, I do not think that the historical Hume held such a view. Anyway, if he did then I think that as an 'empiricist' he would have had to change it in light of such major

more dynamic, pluralistic, and also critically reflective. Philosophers and ordinary persons alike discuss and disagree about the substantive content of particular moral norms, as well as about their justification.[44] If one is interested in Humeanism as a theoretical or methodological approach to explanatory moral theory for modern societies, the 'error theory' objection leads nowhere. Such an explanation obviously does not apply to a moral practice, which is phenomenally very different from the one described in the assumption.

8. Concluding Remark

My conclusion is that one cannot be a Humean and a contractarian. In light of the discussion of Hume's positive alternative to the contractarian ideal of the sources of the normative force of obligations, and of his reconstruction of normative rule-following, it seems appropriate to add two brief remarks. First, concerning the attempt to give a general theoretical account of normative rule-following, there is, I think, a suppressed premise in all of the explanatory approaches touched upon above, which needs to be discussed more openly. That is the implicit *a priori* assumption that there is exactly one right explanatory account of normative rule-following which holds for all persons of all temperaments, and for all moral problems under all circumstances and in all moral environments. I think such an assumption is unwarranted. If the general thesis of *section 6* is convincing, that moral practice is itself critical and self-reflective, then agents can have multiple reasons for following normative rules, and some of these even need not be mutually exclusive. Second, and with respect to the difference between contractarian and Humean accounts of the normative force of obligations, I have presented both accounts in light of the merits and shortcomings that they have, when contrasted from a systematic point of view. This article has primarily tried to identify the crucial systematic concerns, and thus hopefully can indicate a direction of improved analysis that connects the underlying philosophical problems with relevant debates in contemporary philosophy.

References

Baier, K. (1965), *The Moral Point of View. A Rational Basis for Ethics*, New York: Random House.

Chwaszcza, C. (2012), "The Seat of Sovereignty: Hobbes on the Artificial Person of the Commonwealth or State", *Hobbes Studies* 25, 123–142.

---

events as the outcome of the American War of Independence and the French Revolution. Since Hume died in 1776 he did not live long enough to see either event.

[44] A moral theorist like Kurt Baier would argue that a society where this question "Such and such is the rule, but is it right?" has no place, does not even have a morality.

Dworkin, R. (1975), "The Original Position", in: *Reading Rawls*, ed. by N. Daniels, Oxford: Blackwell, 16–53.

Haakonssen, K. (1981), *The Science of a Legislator. The Natural Jurisprudence of David Hume and Adam Smith*, Cambridge: Cambridge University Press.

Hobbes, T. (1994[1651]), *Leviathan*, ed. by E. Curley, Indianapolis: Hackett.

— (1966[1656]), *Elements of Philosophy. The First Section Concerning Body*, *The English Works of Thomas Hobbes* vol. I, ed. by Sir W. Molesworth, London: John Bohn; repr. Aalen: Scientia Verlag.

Hume, D. (1978[1740]), *A Treatise of Human Nature*, ed. by L. A. Selby-Bigge/H. P. Nidditch, Oxford: Clarendon Press.

— (1975[1777]), "An Enquiry Concerning the Principles of Morals", in: *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. by L. A. Selby-Bigge/H. P. Nidditch, Oxford: Clarendon Press, 167–323.

— (1985a), "Of the First Principles of Government", in: *Essays. Moral, Political, and Literary*, ed. by E. F. Miller, Indianopolis: Liberty Fund, 32–36.

— (1985b), "Of the Original Contract", in: *Essays. Moral, Political, and Literary*, ed. by E. F. Miller, Indianopolis: Liberty Fund, 465–487.

Kant, I. (1996[1793]), "On the Common Saying: That May Be Correct in Theory, but It Is of No Use in Practice", in: *Practical Philosophy*, ed. by M. J. Gregor, Cambridge: CUP, 273–310.

— (1996[1797]), "The Metaphysics of Morals", in: *Practical Philosophy*, ed. by M. J. Gregor, Cambridge: CUP, 353–604.

Locke, J. (1988[1689]), "Second Treatise", in: *Two Treatises of Government*, ed. by P. Laslett, Cambridge: CUP, 265–428.

Mackie, J. (1977), *Ethics. Inventing Right and Wrong*, Harmondsworth: Penguin Publishers.

Mill, J. St. (1993[1861]), "Utilitarianism", in *Utilitarianism, On Liberty, Considerations of Representative Government*, ed. by G. Williams, London: J.M. Dent, 1–67.

Rawls, J. (1955), "Two Concepts of Rules", *Philosophical Review* 64, 2–32.

— (1958), "Justice as Fairness", *Philosophical Review* 67, 164–194.

— (1971), *A Theory of Justice*, Cambridge/MA: Harvard University Press.

Ross, W. D. (1930), *The Right and the Good*, Oxford: University Press.

Rousseau, J.-J. (1997[1762]), "Of the Social Contract", in: *Of the Social Contract and Later Political Writings*, ed. by V. Gourevitch, Cambridge: CUP, 39–152.

— (1999[1755]), "Discourse on the Origin and Foundations of Inequality among Men", in: *The Discourses and Other Early Political Writings* ed. by V. Gurevitch, Cambridge: Cambridge University Press, 111–188.

Smith, A. (1976[1790]), *The Theory of Moral Sentiments*, ed. by D. D. Raphael/A. L. Macfie, Oxford: Clarendon.

Sugden, R. (1986), *The Economics of Rights, Co-operation and Welfare*, Oxford: Blackwell Publishers.

— (2009), "Can a Humean Be a Contractarian?", *RMM* 0, 11–23.

Taylor, M. (1987), *The Possibility of Cooperation*, Cambridge: Cambridge University Press.