

VI. HUME'S THEORY OF RIGHT AND WRONG ACTIONS¹

1. *On attributing a theory to Hume*

1.1 *The problem.* The title of this chapter is contentious, for it is not clear that Hume had systematic views concerning right and wrong actions. Certainly he does not make such views fully explicit, and problems relating to them are distant from his usual thoughts about morality. His primary subject is not virtuous actions, let alone right actions understood in a way that allows an action to be right quite regardless of the motive behind it.² His primary concern, being virtuous qualities of mind, is twice removed from matters of right actions. He thinks about morality usually from the perspective of an appraising, ultimately approving or disapproving, spectator and observer, rather from that of a deliberating, perhaps troubled, agent who would decide what to do, what it would be *right* for him to do.³ Perhaps Hume was not very interested in theoretical problems concerning deliberating agents because he did not himself find deliberation and judicious action particularly difficult. Some evidence for his goodness as a person can be gathered from the assessment of his friend's character with which Adam Smith concluded a letter to William Strahan on the occasion of Hume's death:

"Upon the whole, I have always considered him, both in his lifetime and since his death, as approaching as nearly the idea of a perfectly wise and virtuous man, as perhaps the nature of human frailty will permit."⁴

Whatever the explanation, the fact is that there is little in Hume's writings that is explicitly concerned with right actions, and with problems of determining in particular cases which actions would be right.⁵ But there is much that

¹This chapter comes with revisions and amendments from (Sobel 1996).

²*Cf.*: "The last temptation is the greatest treason, to do the right deed for the wrong reason." T. S. Eliot, *Murder in the Cathedral*.

³To call attention to Hume's focus in moral philosophy one might, with a nod to Alasdair MacIntyre's *After Virtue*, title a work on Hume's moral philosophy "Before Duty." I note that amazingly, and in dramatic contrast with Frank Snare (see note 5 below), MacIntyre himself groups Hume with philosophers who attend to rules and principles of right first in their moral philosophizing. He groups Hume with philosophers who do *not*, as Hume certainly does, "attend to *virtues*... first" (MacIntyre 1984, p. 119).

⁴"Letter from Adam Smith, L.L.D., to William Strahan, Esq., Kirkcaldy, Fifeshire, Nov. 9, 1776." "In this last sentence Adam Smith was deliberately echoing the last sentence of Plato's *Phaedo*." (*For this reference to Plato, I am indebted to D. D. Raphael...)." (Mossner 1980, p. 605.)

⁵Frank Snare says that Hume "does not recognise a **distinct**... class of rightness judgments" – a class of judgments distinct from virtuous judgments (Snare 1991, p. 28, bold emphasis added). This charge goes beyond the evidence which, with one exception, is only that Hume rarely discusses right and wrong, as distinct from virtuous and vicious, actions. The exception is that Hume does say "[a]ctions themselves [without regard to characters of which they are signs]...are never considered in morality" (T132 [575]). The sentiment of this text is nowhere expressed in works subsequent to the *Treatise*, which book is described

(continued...)

bears implicitly on questions of right action and moral deliberation. There is so much that bears on these things, and the implications are sometimes so obvious, that it is neither unusual nor in my view objectionable for commentators to write as if Hume was concerned directly and explicitly much of the time with actions, and with right and wrong.⁶ For Hume provides a very detailed account of the virtues and thus by easy implication of the thoroughly virtuous person. And so, given that right actions must be somehow related to thoroughly virtuous persons, he provides by a not so easy implication a theory of right actions. The first thing to settle in order to draw it out, is *how* he would say the ideas of right actions, and of thoroughly virtuous persons are related. He does not explicitly say.

1.2 *An interpretive assumption.* An interpretive assumption that may come first to mind is that Hume would say, if asked, that the ideas of moral rightness and perfect virtuousness are connected 'by definition' thus:

an action A is morally right in a case
if and only if,
were the agent ideally well-informed and possessed of every moral virtue
in the manner of a perfectly virtuous agent in an otherwise similar case, then he might do A.

There can be several actions each of which is morally right in a case. That is why 'might' figures instead of 'would'. That one may feel is in broad outline *the way* of finding in a theory of virtue and the virtues, a theory of actions right and wrong.⁷ It would be what Michael Smith could term an '*example-model*' interpretive assumption (see Smith 1995). Robert Neal Johnson says that that model "fits nicely with a normative theory grounded in ideals of character" (Johnson 1997, p. 620). There are, however, difficulties with 'example-model' theories of moral right and wrong, as there are for such theories of what would be among best, and other than best actions, everything considered, "for perhaps [going back to theories of moral right and wrong] what an agent would do were [he fully

⁵(...continued)
in Hume's Advertisement to the second volume of the posthumous edition of *Essays and Treatises of 1777*, as "that juvenile work, which the Author never acknowledged" [E2]. And this exceptional text is embarrassed by others, several of which will be quoted, in which Hume writes of right and wrong, duty, and obligation in ways that imply that these are functions exclusively of tendencies of actions and practices to have useful consequences, and not at all of motives and characters of agents. Hume, at (T132 [575]) is thinking only of contexts assessment and evaluation to 'love or hate, praise or blame', and not also of contexts of deliberation what to do, and what to advise.

⁶For a case to this point I offer the able commentary of (Harrison 1981), in the preface to which, Book III of Hume's *Treatise* is said to contain "Hume's account of our knowledge of right and wrong" (p. vii).

⁷"[I]f the question is, 'How can virtue ethics give an [account of right action in such a way as to provide action guidance?]' the answer is easy. Here is its first premise. (P1) An action is right iff it is what a virtuous agent [i.e., one who has, and exercises, certain character traits, namely, the virtues] would characteristically (i.e. acting in character) do in the circumstances." (Hursthouse 1999, p. 28.)

informed and thoroughly virtuous, and in a case otherwise like a case 'in hand']...would, given his actual state of heart and mind, have relatively bad consequences" (Sobel 1982, p. 123).⁸ That, in the view of any theorist who, like Hume, thinks consequences can count, should make for possible differences between what he would do if thoroughly informed and virtuous, and what it would be right for him to do.

With such difficulties in mind, I propose, instead of that 'example-model' interpretive assumption, what Smith might term the '*advice-model*' interpretive assumption that Hume would say that, 'by definition',

an action A is morally right for an agent X in a case C
if and only if

if X were *ideally well-informed* [I] and possessed of every moral virtue in the manner of a perfectly virtuous agent, then he would, so far as dispositions of his *moral virtues* [II] to favour and disfavour kinds of action [III] go, want a person *exactly like* [IV] the person he (X) actually is in C to do in a case exactly like C any action from a set that includes an action exactly like A.

The 'idea' of this 'definition': Let X* be related to X as in this 'definition' of 'morally right' in terms of moral virtues: that is, X* is X himself perfected in certain ways. Pretend that X asks X*, "What, morally speaking, ought I to do in this situation?" The idea is that the *moral advice* X* would give X, would agree exactly with what X*, so far as his moral dispositions to favour and disfavour actions go, would *want* X to do.

[Notes on terms in the definition: [I] an *ideally well-informed* person would have what were for him ideally well-informed probabilities for circumstances of C and possible consequences of his actions in C, but not necessarily *true beliefs* of these. That this stipulation regarding 'ideally well-informed' introduces, under the rubric '*for him* ideally well-informed probabilities', elements of agent-relativity and subjectivity into this virtue-based 'definition' of right action, is an argument for requiring of an ideally well-informed person true beliefs of circumstances and consequences, but not, I think, for interpreting Hume's implicit theory of right action: see Section 4.1 below. [II] I take '*moral virtues*' to be, for Hume, principally the virtues of benevolence and justice, along with several other social virtues including obedience to the state, and 'responsible parentage' that involves a sense of duty to care for one's children, but does *not* include loving them where this love would involve being *partial* to them – see Section 3.3 of the previous chapter. (III) The

⁸The 'hero' of Percy Walker's *The Moviegoer*, for guidance would ask, "What would Fred Astaire do?" Often, I think, the answer would be, "Dance!" and were our hero to dance, he would have danced badly.

moral virtues, I recall from Chapter II, Section 2 above, include not only dispositions to perform certain kinds of action, but dispositions to favour performances of these actions – similarly for dispositions to avoid and to disfavour certain kinds of actions. An honest person, for example, is not only disposed not to cheat, but to be for other persons's not cheating, to want others not to cheat. An honest person thinks from time to time of cheating as *wrong*. Such thinking includes disfavour of cheating, of acts of dishonesty. (IV) Taking a clause from Richard Hare, 'exactly like' is here short for 'exactly like in every universal respect': for example, if John's mother is Mary, another person named 'John' can be 'exactly like' John for purposes of this 'definition' of right action even if his mother is not Mary, as long as she is like Mary in every universal respect.]

An agent, to qualify as a 'measure' of right and good needs to be perfectly moral, or, in other words acceptable to Hume, perfectly "*virtuous*" and "a man of virtue [*par excellence*]" (E284 [314]). He needs to have every *moral* virtue, but he need not have every quality of mind that Hume counts as a 'virtue', and so, though he must be worthy of the highest moral esteem, he can be only moderately admirable and worthy of esteem *tout court*: he can be boring and a bit of a prig, not particularly cleanly, and rather dull-witted. But he will have every moral virtue. He will be properly kind and honest and the rest. It is better to say 'properly' than 'perfectly' here, since he must not be *unlimitedly* just or benevolent. This very model of a man of virtue, this perfectly virtuous man, will according to Hume's view not be *extreme* in his dedication to justice, nor in his benevolence to his fellow human beings; for when his dispositions to justice and benevolence conflict, as Hume says they may often do, sometimes one of these cardinal virtues should be displayed in the choice and action, and sometimes the other.

The 'advice-model' interpretive assumption that I am making is not as simple as the 'example-model' interpretive assumption. But its details, in so far as they are different, provide for a more plausible and defensible theory of right than do those of the 'example-model' assumption. This 'advice-model' interpretive assumption grounds right and wrong in "ideals of character" (Johnson 1997, p. 620) in a manner that is in one way less 'nice' since it is more complicated, but that is in this other of plausibility, that counts for more, 'nicer'. In the sequel, however – that is, in the theory of right and wrong that I project for Hume – no stated details turn on the difference between these two interpretive models. I have made the distinction and attributed an 'advice-model' to Hume pursuant to my policy of generous interpretation: in my view the 'advice-analysis' is a better analysis.

1.1.3 What kind of proposed ‘definition’ or ‘analysis’ would that interpretive assumption be? It would be a ‘summary analysis’ in Michael Smith’s terms (Section 4.3 of Chapter II above). It would not be a naturalistic, reductive, summary analysis, for the definition posited in it does not run entirely in ‘natural terms’ since it would define moral words for actions, ‘right’ and ‘wrong’, partly in terms of a moral word for qualities of mind, ‘virtue’. Jonas Olson has observed that it would *say* what it is *to be* right and wrong, *not* what it is that *makes* actions right and wrong. It would say that to be right or wrong is to have the power to elicit favour or disfavour of a kind from a certain ideal person on a certain view. That, however, entails an analysis of the relations of ‘makes right’ and ‘makes wrong’.⁹ What makes actions right and wrong according to this posited analysis of rightness and wrongness is what about right and wrong actions gives them those powers. The principle of action stated in Section 2.1 below would say what, in general, gives actions these powers, and thus what makes them right and wrong.

1.2 *Texts on the practice of morality.* The general lines of Hume's implicit theories of right actions and ideal moral deliberation, and the main problems and complications it runs into, can be gathered from texts in the following two subsections.¹⁰ These texts are among the few that are explicitly concerned with assessments of actions in terms other than the characters of their agents and motives.¹¹

1.2.1 *The purpose of virtue, and its demands.* Utility and the greater happiness are, we are told, objects of constant and indeed sole concern in moral evaluations of actions.

⁹These relations ‘want’ to be explained. Thinking of the ‘objectivist views’ that he considers parts of the common sense of morality, and of objectivist theories of morality such as those of the Intuitionists, John Mackie complains: “What is the connection between the natural fact that an action is a piece of deliberate cruelty...and the moral fact that it is wrong? It cannot be...logical or semantical necessity....The wrongness must somehow be ‘consequential’ or ‘supervenient’: [in the example] it is wrong because it is a piece of deliberate cruelty. But just what *in the world* is signified by this ‘because’?” (Mackie 1977, p. 41.)

¹⁰Hursthouse writes of the general lines and problems of the implications of ‘virtue ethics’ for right and wrong actions, meaning, however, by ‘virtue ethics’ *neo-Aristotelian* virtue ethics. She realizes that there is “a lingering view that [ethics] is, or ought to be, more codifiable than virtue ethics makes it out to be” (Hursthouse 1999, p. 41) and offers to dispel it. Hume’s virtue ethics has, according to my reading of its implications for right and wrong, invites no such view: I think no one will say of the theory of right and wrong that I pin on it, that there is *more structure* to determinations of right and wrong than it says.

¹¹The following text concerns deliberation, but implies nothing for general lines of Hume’s theories of right action and ideal moral deliberation: “When a man...deliberates concerning his own conduct (as, whether he had better, in a particular emergence, assist a brother or a benefactor), he must consider these separate relations, with all the circumstances and situations of the person....in moral deliberation we must be acquainted beforehand with all the objects, and their relations....” (E265-6 [289-90]).

"[I]t appears to be [a] matter of fact that the circumstance of *utility*...is **constantly appealed** to in all moral decisions concerning the merit and demerit of actions...and...that it is a foundation of the chief part of morals, which has a reference to mankind and our fellow creatures." (E221 [231], bold emphasis added)

"[Virtue] declares that **her sole purpose** is **to make** her votaries **and all mankind**, during every instant of their existence, if possible, cheerful and **happy**....**The sole trouble which she demands**, is that of **just calculation**, and **a steady preference of the greater happiness**." (E258 [279], bold emphasis added)

These last words, especially the word 'sole', can suggest that Hume thinks that the trouble virtue demands is *merely* 'calculating' and comparing tendencies of actions to promote happiness, as if virtue does not lay heavy deliberative burdens on her votaries, the virtuous. But Hume is concerned, he is very concerned, to bring out that this suggestion is in two ways far from the truth. First, 'the trouble of calculations' of the tendencies of choosable alternative actions to promote happiness, the trouble of reaching 'nice views' of 'balances', can be *considerable*. And second, 'calculations' to do with the general happiness involved in deliberations of a virtuous person can be much less *straightforward*, can be far *more intricate and involved*, than are calculations or assessments of tendencies of available alternative actions for the general happiness.

1.2.2 *The main difficulties virtue makes for an agent*

"One principal foundation of moral praise being supposed to lie in the usefulness of any quality or **action**, it is evident that *reason* must enter for a considerable share in all decisions of [of praise and censure]; since nothing but that faculty can instruct us in the tendency of qualities and **actions**, and point out their beneficial consequences to society and to their possessor. In many cases this is an affair liable to great controversy: doubts may arise; opposite interests may occur; and a preference must be given to one side from **very nice views**, and a **small overbalance** of utility. **This is particularly remarkable in questions with regard to justice; as is, indeed, natural to suppose, from that species of utility which attends this virtue**. Were every single instance of justice, like that of benevolence, useful to society; this would be a more simple state of the case, and seldom liable to great controversy. But as single instances of justice are often pernicious in their first and immediate tendency [and indeed – as Hume sometimes indicates – in their total tendencies], and as the advantage to society results only from the observance of the general rule, and

from the concurrence and combination of several persons in the same equitable conduct; the case here becomes **more intricate and involved.**" (E262 [285-6], bold emphasis added.)

One can imagine!!

Hume writes sometimes, as here, of the 'first and immediate tendencies of acts of justice', sometimes of these acts 'were they to stand alone', sometimes of them 'considered apart' (T66 [498]), and sometimes of them 'considered in themselves' (T135 [579]). He says several things about ways in which tendencies of systems of actions can diverge from tendencies of their component actions, including that acts of justice 'considered in themselves' can be contrary both to the public and private interests, though they make general schemes or systems of actions that are essential to both. Other words, that are not his, for this point, that is, for Hume's main and best point here, are that acts of justice '*considered singly or distributively*' can be contrary to both public and private interests, whereas '*considered together or collectively*' they are essential to interests both public and private. In the next chapter I go into delicate matters concerning the possible 'logics' of the 'species of utility that attends justice' of which Hume speaks, and come back to this topic when discussing 'consequences of actions' in Chapter VIII below. In the present chapter, without further discussion, and without full explanations, I assume for this species of utility a radical, considered-singly/considered-together, 'logic' that his best words for it quoted in Section 12.3.3 of the previous chapter.

1.3 *Towards a theory*

1.3.1 In his essay "Of the Original Contract" (1752, the year following publication of his *Second Enquiry*), Hume distinguishes between two kinds of 'duties' which suggests a two-part utilitarian theory.

"All *moral* duties may be divided into two kinds. The *first* are those to which men are impelled by a natural instinct or immediate propensity which operates on them, independent of ideas of obligation and all views either to public or private utility. Of this nature are love of children, gratitude to benefactors, pity to the unfortunate. **When we reflect on the advantage which results to society from such humane instincts, we pay them the just tribute of moral approbation.** But the person actuated by them feels their power and influence antecedent to any such reflection.

"The *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 [that] *justice*, or a regard to property of others, *fidelity*, or the observance of promises, become obligatory and acquire an authority over mankind." (Hume 1948, pp. 366-7, bold emphasis added.)

The first kind of 'duty' are 'moral' Hume says given their 'advantages to society', that is, in terms from the previous chapter, given their 'general utility'. This suggests that the first kind of duty can be subsumed under 'appeals to benevolence' that do not presuppose institutions of rules and conventions, with the second kind being made up of demands of justice that do presuppose institutions of rules and conventions (for example, that define promises and distinguish them from mere statements of intentions, and that define property and distinguish it from mere possessions)^{12, 13}. The *problem* left by this division is what happens when duties of these two kinds *conflict*.

1.3.2 *Usually, but not always, when benevolence and justice conflict, justice prevails.* That rules of justice *should* 'come first' is implied by the 'standing' of justice amongst the virtues – "no moral excellence is more highly esteemed" (E200 [203-4]) – and by "[t]he necessity of justice to the support of society" (E200 [203]), and the great happiness for which society is itself necessary (*cf.*, (E188 [188])). Also, rules of justice *count*, only if they *sometimes* come first, since the duty of generalized benevolence, the duty advantage on balance all affected, has application in every case. But the priority due to the 'rules of justice' is not – indeed, given the basis for the virtue of justice, which is the general utility of its widespread incidence – it *cannot* be absolute.

"Does any one scruple, in extraordinary cases, to violate all regard to private property of individuals, and sacrifice to public interest a distinction, which had been established for the sake of that interest?" (E196 [195])

Sometimes benevolence wins, it wins in *extraordinary* cases, and so *ordinarily*, that is *usually*, justice that prevails.

[On a related, but somewhat problematic note, we have that,

¹²In a pre-society 'state of nature' an apple can be possessed but not owned, so that when I put it down it is 'up for grabs', and that I will give half of it to you can be stated by me as an intention but not promised, so that my changing my mind is, in itself, not grounds for complaint.

¹³The division that Hume draws between two kinds of 'moral duties' is troubled by 'debts of gratitude' and 'parental duties' which, though not presupposing 'social artifice', are supported by 'original instincts', and which do not when neglected bring down society. Love of one's children and wanting to care for them is common enough to make neglect of *duties* to care for them relatively rare and of little consequence.

"Nothing less than the most extreme necessity, it is confessed, can justify individuals in a breach of promise, or an invasion of the properties of others." (E203 [206])

Evidently 'extreme necessity' *can* justify breaches of rules of justice. Though when *it* does it seems that not benevolence, but interest prevails, albeit only *extreme* interest. The problem is why how *much* keeping a promise of staying off another's land would cost a would-be promise-keeper and trespassers, should matter to his obligation to honour that promise and to respect the other's property. There is more on this small puzzle below, in Section 3.2.4.]

1.3.3 *What matter, pace Immanuel Kant, are duties defined by actually in place in one's society useful rules.* I attribute to Hume a mixed act/rule form of utilitarian theory. On its rule-side, which establishes the duties of justice, what matter are rules that are *actually in place*. It is required that these be useful, but to matter, *they need not be ideal*, either theoretically, or even practically. On the first point, that they need not be theoretically ideal for their subjects, we have that:

"A rule, which in speculation may seem the most advantageous to society, may yet be found, in practice, totally pernicious and destructive." (E193 [193])

That Hume's implicit theory of right and wrong accords precedence only to actual rules of 'up and running useful practices' is evident in many passages. It can be gathered, for example, from Hume's report of the "suspension of justice among...warring parties" (E188 [187]), and from what he says must be the conduct of a virtuous man whose fate it is,

"to fall into the society of ruffians, remote from the protection of laws and government...[H]is particular regard to justice being no longer of use to his own safety or that of others, he must consult the dictates of self-preservation alone." (*Ibid.*)

These lines express Hume's view of the irrelevance, to an individual's right actions, of 'rules of justice' that are being generally ignored. Though the *reason* he gives for this irrelevance is not consistent with Hume's view of the species of utility of justice according to which particular regards to justice are *often* useless, indeed pernicious, in every way, even when, since there is general regard to justice that is highly advantageous for all and each, justice requires particular regards. Hume should have said that the reason for the irrelevance of a person's justice when he has fallen in with ruffians is not that his particular regard to its demands would not in these circumstances be useful, but that in these circumstances there is not general regard to any rules for it. There is, in the case envisioned, "remote[ness]

from "...laws and government" and society in which rules of justice are actually in place and widely honoured. 'Here' in this "society [*sic*] of ruffians" there is "disregard for equity...contempt of order...**a total dissolution of society**" (E188 [187], bold emphasis added). Hume writes of a person who has fallen into a most unlikely 'society' of *ruffians* amongst whom 'there is no honour'. This would not be a society of *bandits* amongst whom there *is* honour, and there are rules.¹⁴ The reason why this unfortunate man need regard justice no longer, is that he has come to a place where *there is no justice*, where there are no rules for justice that are generally honoured, and the honouring of which is generally demanded as right and proper. That, according to Hume, is what makes the difference here.¹⁵

The question of justice is always the question what *is* required by useful rules of it, *if any are in place*. The question is never, for Hume, what *would be* required by rules of justice, useful or ideal. Peace, therefore, to Kant, who is of a very different view.

"[T]hough he scrupulously follow [the rules of morality, he] cannot for that reason expect every other [or even *any* other] rational being to be true to [them]...Still the law: Act according to the maxims of a universally legislative member of a merely potential realm of ends, remains in full force." (*Foundations* 438-9.)

1.3.4 The theory of right and wrong that I believe is implicit in Hume's discussion of the moral virtues, while complicated and intricate in its treatment of 'rules of justice', as his remarks on the species of utility of justice require that it be, is throughout *purely* utilitarian. It does not, for reasons indicated in Appendix A to this chapter, include 'contractarian factors' of private utility that Hume's remarks on conditions under which 'rules of justice' take place and are established have seemed to some readers to imply.

1.3.5 *On varieties of utilitarian theories of right actions*. *Pure utilitarianisms* make tests of *general utility*, variously applied, decisive for right actions. *Act-utilitarianisms* apply tests of general utility directly only to acts. For a

¹⁴*Cf.*: "Robbers and pirates, it has often been remarked, could not maintain their pernicious confederacy, did they not establish a new distributive justice among themselves, and recall those laws of equity, which they have violated with the rest of mankind." (E205 [209].)

¹⁵But what if a virtuous man falls into a society of bandits amongst whom there is honour, and he is not one of them? Then, I assume, it would be understood all around that its rules did not *apply* to him. That could be evident from his being shackled and guarded, and his being expected not to keep his word not to cooperate with the authorities of his own society in the prosecution of these bandits, should he be released to return to it.

simple act-utilitarianism, there is the principle that an action is right if and only if it is amongst those open to its agent at its time that would produce most happiness. *Rule-utilitarianisms* apply tests of general utility directly also, or only, to rules (more exactly, to general conformities to, or acceptances of, rules), and either never, or not always, also to particular actions. *Ideal-rule utilitarianisms* apply tests of general utility directly also, or only, to rules to find *best possible rules*, without regard to rules, if any, that are actually in place, and are being followed by most people. For one theory of this type, we have the principle that an action is right if and only if it accords with those rules (for members of a community that includes its agent – this may be ‘the community of humankind’, or some smaller accidental community) universal conformity to which would produce most happiness (in this community). *Actual-rule utilitarianisms* apply tests of general utility directly also, or only, to rules, to verify the general usefulness, variously specified, of actually established and generally adhered to rules. It is a pure utilitarian theory of this last type that I ‘find’ in Hume’s work on morals.

2. *A sometimes actual-rule, sometimes straight-act, pure utilitarian principle for Hume*

2.1 *The principle, and a decision procedure suited to it.* The theory that I attribute to Hume is summarized in the following principle of right action.

An action A is ‘right’¹⁶ in a case C

if and only if:

EITHER

(*for one thing*) there is at least one rule R that ‘applies’ in case C such that, (i), R is a ‘contributing member’ of an ‘established’ and ‘socially useful’ system of ‘rules of justice’ for the time and place of C, and, (ii), C is not an ‘extraordinary case’ in relation to R,

(*for a second thing*) action A conforms to every rule (i.e., does not violate any rule) that satisfies (i) and (ii), and

¹⁶Scare-quotes here and in the title to this chapter in deference to Hume, who uses the words ‘right and wrong’ in relation to justice and its subdivisions (T67 [498]: “as to the *moral* obligation [to justice], or the sentiment of right and wrong...), and of rules for practices similar to those of property and communication (E206 [211]: war and sportive war “beget...a standard of right and wrong...”), but never in relation to benevolence and its subdivisions. The words ‘right and wrong’ occur several times in Book III, Part I of the *Treatise* and in Section I of the *Second Enquiry* when virtues are not in view.

(for a third thing), of actions that conform to every rule that satisfies (i) and (ii), action A would, in terms of the utilities and ideally well-informed likelihoods of its agent's for possible consequences, be at least as attractive from the standpoint of general happiness and utility as would be any other action that is open to the agent that satisfies (i) and (ii);

OR

(for one thing) there is no rule R such that R satisfies both of conditions (i) and (ii), and

(for a second thing) action A would, in terms of the utilities and ideally well-informed likelihoods of its agent for possible consequences, be at least as attractive from the standpoint of general happiness and utility as would be any other action that is open to the agent.

For a principle of right *and wrong*, we may add that **an action is wrong in a case** if and only if it is not right in this case.

Comments on terms in this principle

A rule *applies* in a case if and only if it is possible to violate it in this case.¹⁷ A rule is a '*contributing member*' of an established and socially useful system of rules, if its deletion would detract from that usefulness. A *socially useful* system of rules is one that makes affected lives better than they would be 'in the jungle', where there is no order, and no society. A 'socially useful' system in this sense is, one might say, a *minimally* social useful system: it is a system sufficient to some order and society within which persons can interact for more happiness than is possible 'in the jungle'. In the principle, and in what follows,

'rule of justice'

complete with single stipulated-sense quotes included, is short for *rule of the type of a rule of justice*. To illustrate, rules of allegiance to governments and of obedience to governmental authorities, as well as rules for chastity and modesty (not even a 'glimpse of stocking'), and against adultery, while not rules of justice *proper* for Hume, are said to have the characteristics that I say imply, for rules of justice proper, preeminent positions in moral deliberations. Rules of justice proper include not only rules of property, but also certain rules of discourse and of commerce. I assume that Hume thinks of veracity and fidelity to promises, but *not* also of allegiance, chastity and modesty, and

¹⁷The rule that calls for sharing your lunch if you are eating with someone who has not brought enough, does not apply in a case in which you are eating alone.

fidelity to the marriage bed, as "subdivisions of justice" (E277 [305]). Additional comments on 'rules of justice' in the principle, as well as all comments on '*extraordinary cases*' are coming in Section 3. For a theory of *established social rules*, one cannot, I think do better than the theory elaborated by H. L. Hart in *The Concept of Law*, Oxford: Clarendon Press, 1961.

2.2 Complementing this rather complicated principle of right action, and serving to bring out its sense, are the following instructions for a *moral decision procedure* suited to it.

To decide what it is morally required that one do: First, assuming that you are satisfied that the system of established 'rules of justice' for your time and place is socially useful, identify contributing members of it that apply to your case, and under which it is not an 'extraordinary case'. Next, if there are any such rules, then – having settled in your mind (as best you can in the time available that as far as you can see the issue before you deserves to have expended on it) utilities and likelihoods of relevant possible consequences¹⁸ – apply the test of general utility, in turn, to each action that conforms to these rules, and do one these that scores highest on the test. If, on the other hand, there are no such rules, then – having settled in your mind (as best you can in the time available that as far as you can see the issue before you deserves to have expended on it) utilities and likelihoods of relevant possible consequences¹⁶ – apply the test of general utility to each action open to you, and do one that scores highest on these tests.

The 'spirit' of this principle of right action, and the decision procedure suited to it, that I say are implicit in Hume's books on morals, has the following expression:

Assuming that you are satisfied that the system of 'rules of justice' for your time and place is socially useful, which it almost certainly is: as best you can in the time available that as far as you can see the issue merits to have expended on it, settle relevant likelihoods and utilities, *and promote in your actions general utility and the happiness of mankind, subject to the constraint to*

¹⁸There is here 'a can of worms' for practical deliberation that would take into consideration consequences, for (to open this can) there is the question, How much time to spend on the question, How much time does the issue of the case deserve?, *and* there is the question, How much time to spend on *that* question?, *and so on* to higher and higher order time-limit questions *ad infinitum!*

conform, except in 'extraordinary circumstances', to established 'rules of justice' the general observance of which is contributing to that social usefulness.

The prescription is for promotion of the general happiness, but not *unfettered* promotion. The prescription is to promote the general happiness *after* one has taken care of the business of justice and its rules, which business is in its description, as it can be in practice, *complicated*, and in other than extraordinary circumstances prior.

2.3 *That disputes in morals all go to public utility*

"In all determinations of morality, this circumstance of public utility is ever principally in view; and wherever disputes arise, either in philosophy or in common life, concerning the bounds of duty, the question cannot, by any means, be decided with greater certainty, than by ascertaining, on any side, the true interests of mankind." (E183 [180])

'Disputes in *philosophy*' referenced here are, I believe, concerned not with 'bounds of duty,' but with 'bounds of virtue'. Such disputes are according to Hume rare.

"The quick sensibility, which, on this head, is so universal among mankind, gives a philosopher sufficient assuranceThe very nature of language guides us almost infallibly in forming a judgment of this nature...." (E178-9 [174]).

Hume says, mistakenly, that when controversy occurs it is easily resolved. For example: "Giving alms to common beggars is naturally praised... ['but on only a *little* reflection' I hear Hume saying] we regard that species of charity rather as a weakness than a virtue" (E183 [180]).

'Disputes in *common life*' referenced are, as Hume says, to 'bounds of duty' and are over right and wrong in particular cases, and in kinds of cases. Such disputes are in Hume's view not so rare. Nor are they always easily settled.

"One principal foundation of moral praise...[lies] in the usefulness of any quality or action....In many cases this is an affair liable to great controversy....[particularly so] in questions with regard to justice...." (E262 [285], quoted in full in Section 1.2 above).

2.4 According to my reading, "the supreme law" which is "public safety and interest" (E195 [196]) would in ideal deliberation never be applied without further ado to decide what is right. An ideal deliberator would always check

first whether established rules of a certain sort are relevant to the case. If some are, he would apply a test of utility not to available actions, but to those relevant rules. It would be applied to verify that they are contributing members of socially useful systems of up-and-running rules, which is *not* to say that they are contributing members of best possible systems; it *is* to say only that they are contributing members of an established system of rule establishment of which has, from the standpoint of general utility, made conditions for living better than they would be were we to fall back into a jungle of no 'rules of justice'. However, though in *no* case would tests of utility be applied to actions 'without prior ado', and though in *some* cases they would not be applied *first* to actions before they were applied to anything else; tests of utility would in *every* be applied *eventually* directly to actions. In every case in which they are not applied first to actions, because there are established rules in the picture of the relevant type, they are applied *eventually*, in connection with condition (ii) of the principle to conformities to rules that satisfy condition (i), and possibly again in order to choose among actions that would conform to all rules that satisfy conditions (i) and (ii). It is a complicated utilitarian theory of right actions. It is complicated because of the peculiarity of "the species of utility that attends" justice (E262 [285]) that it accommodates, and the qualified priority that it accords to established socially useful 'rules of justice'. However, that the picture of ideal moral deliberation implied is complicated, does not argue *in the least* that it is not a true picture.

3. *Ancillary theories*

To complete the theory I attribute to Hume, sub-theories are needed, theories or accounts of the intended sense of some terms in the principle of right action, one of which has been glossed in Section 2.1. Now come additional remarks concerning *rules of justice*, and a theory of *extraordinary cases*.

3.1 '*Rules of justice*,' that is, *rules of the type of rules of justice*

3.1.1 *What is distinctive about these rules.* I assume that among these rules are either the following, or versions of them in which conditions are spelled out, and exceptions stated:

Keep your promises.	Do not trespass.
Do not steal.	Tell the truth.

In contrast, rules and taboos such as,

Refrain from assaults.	Do not kill.
------------------------	--------------

Don't be cruel.¹⁹

are *not* 'of the type of a rule of justice'.

Distinctive of 'rules of justice' is that they have that 'species of utility which attends justice', or something *like* this species of utility. Of 'rules of justice' that have that species of utility 'full bore': *General* observances of rules of this type in a society are of utility not only on balance for this society,[*] that is, *for all*, but – with the exception of 'self-sufficient supermen' (if such human beings can be)²⁰, and persons who are exploited for the good of all considered together[**] – also *for each* though *this* is not a part of the *idea* of that species of utility or of a 'rule of justice'. Hume casts general observances of rules of justice as social necessities,²¹ as necessities for *humanity*.²² And yet *individual observances* considered singly are often not of social utility, let alone being social necessities, and are often not of private utility either: they are often 'pernicious in every way'. If such a rule were observed only when particular individual observances of it were recommended by their individual social utilities or private utilities, much, perhaps all, of the social utility and the private utilities of its general observance would be lost. *Cf.*:

"[H]owever [much] single acts of justice may be contrary either to public or private interest, it is certain that the whole plan or scheme is highly conducive, or indeed absolutely requisite, both to the support of society and the well-being of every individual....[W]hatever may be the consequence of a single act ... yet the whole system ...is infinitely advantageous to the whole and to every part" (T66 [497-8]).

¹⁹'To a heart that's true' Elvis Presley sang.

²⁰They would have to be born with the powers of adults, in order not to be beholden to general observance of the inter-generational 'rules of justice' of responsible parentage supposing that parental love was strangely lacking. (Of course, Paul Grice and Peter Strawson have taught, it is, *pace*, Willard Van Orman Quine, *analytic* that though born with the powers of adults, they would not be born adults.)

²¹I am not making this a defining feature of 'rules of justice' in the principle of Section 2.1. What is distinctive of them, I am saying, is *exactly* that they have that 'species of utility that attends justice'. It is possible that an established, generally useful rule should have that utility, though it is not part of a system of rules that are necessary for society. A more complicated principle, closer perhaps to the best principle that can be projected from Hume's texts, would modify (i) somehow to confine it to systems of rules *S* such that *S* is a member of a set *S* of systems of rules such that it is necessary for society at the time and place of *C* that some system of rules in *S* be established.

²²Hume traces "the virtue of chastity or fidelity to the marriage bed" (mistaking for one virtue what would be two) to "[t]he long and helpless infancy of man [that] requires the **combination** of parents for the subsistence of their young" (E203 [206-7]; bold emphasis added). These 'rules of the type of rules of justice' are for general happiness of human kind through the ages.

"Did all his views terminate in the consequences of each act of his own, his benevolence and humanity, as well as his self-love, might often prescribe to him measures of conduct very different from those which are agreeable to the strict rules of right and justice." (E278 [306]).

[*] Which is to say that their general observance has 'general utility' in the sense of Section 1.1 of the previous chapter: "a quality of mind has 'general utility' if the balance, when from the sum of its agreeable effects is subtracted the sum of their disagreeable opposites, is positive." Something can have 'general utility', and be advantageous for persons affected *considered together*, though it is not advantageous to each person affected.

[**] Will an established 'rule of justice' *extend* to persons its general observance does not, for one reason or another, advantage? Will such persons have either rights or duties under it? No. This is because 'rules of justice' *that take place* and are *established* extend only to persons mutually advantaged by their establishment. While it is not necessary that general observance of a 'rule of justice' be advantageous to each person *affected by it*, it is necessary that every established 'rule of justice' should be advantageous to each person *to whom it extends*.

Cf.: "Were there a species...intermingled with men, which, though rational, were possessed of such inferior strength...they they were incapable of all resistance, and could never, upon the highest provocation, make us feel the effects of their resentment;[#] the necessary consequence, I think, is that we should be bound by the laws of humanity to give gentle usage..., but should not, properly speaking, lie under any restraint of justice with regard to them, nor could they possess any right...exclusive of such arbitrary lords...[A]s no inconvenience ever results from the exercise of power, so firmly established in nature, the restraints of justice and property, being totally *useless, would never have place* in so unequal a confederacy." (E191 [190-1], bold emphasis added.)

Section 5.2 worries consequences for institutions of slavery of the principle of right and wrong I attribute to Hume. Appendix A elaborates on the theme of 'conditions under which rules of justice take place'.

[#] Did Hume think there were such beings, such 'natural slaves'? No. "The great superiority of civilized Europeans above barbarous Indians, **tempted us to imagine** [they were]...and made us throw off all restraints of justice, and even humanity, in our treatment of them. In many nations, the female sex are reduced to like slavery...yet...women are commonly able to break the confederacy, and share...in all the rights and privileges [and duties] of society." (E191 [191], bold emphasis added). Hume did not, as Aristotle did, believe in 'natural slaves' and the analogous natural inferiority of women. [*Politics* I,4,1254b13-8; cited in Section 4.4 of Chapter IV (Sobel 2005c)].

Included amongst 'rules of justice' of the principle of Section 2.1 are rules that have only '*something like*' that species of utility that attends rules of justice proper. Specifically, the umbrella term 'rules of justice' covers rules whose general observances have social utility, and are good for all (and generally for each, other than supermen and exploited persons, if any – but again this goodness for each is not *required* or part of the idea of a 'rule of justice'), though their individual observances are often lacking in *private* utility, and are indeed often of significant private disutility. Such rules should have priority in the principle, even if individual observances of them are never lacking

in *social or public* utility. Constraints of justice and the like are in Hume's view (E185ff [183ff]) *especially* for the common conditions of limited resources *and limited generosity*.

3.1.2 '*Rules of justice*', and '*generalization arguments*'. A mark of a rule of the type of justice is the special relevance to its injunctions of the question, "But what if *everyone* did that?" Against indifference to property distinctions, and in general to requirements of justice, we bring not the disutility of particular transgressions, but of general practices.

"What must become of the world, if such practices prevail? How could society subsist under such disorders?" (E201 [203], original italics)

'Rules of justice' proper are seen as not merely useful, but necessary for society. They are seen as essential to society and a social order in which people can interact and thrive, as necessary to *minimal* happiness, and as providing the framework for much more. Without society and social order there can be, Hume would agree,

"no place for industry...no culture of the earth...and the life of man [is] solitary, poor, nasty, brutish, and short." (Thomas Hobbes, *Leviathan*, Chapter 13, "Of the Natural Condition of Mankind as concerning their Felicity and Misery.")

With society, 'everything is possible'.

Socrates, when defending the rule to honour court decrees – a rule of the type of a rule of justice under the general rule of allegiance, Hume would say – has the Laws ask, without waiting for an answer, whether he thinks it is "possible for a city not to be destroyed if the verdicts of its courts... are [regularly] nullified and set at naught by private individuals" (*Crito* 50b, G. M. A. Grube translation). Such questions – such *generalization arguments* – have natural relevance to promises, truth, and the law; they seem just the arguments to use, against, for example, breaking promises, lying, stealing, and breaking the law, especially in the all too frequent cases in which it seems that no harm would be done by a particular broken promise, lie, theft, or breach of the law, and considerable harm would be *avoided*.

'Generalization questions', if offered as rhetorical arguments against acts of assault, killing, and gratuitous cruelty, can seem odd. There are always better and simpler things to say against acts of cruelty, assaults, and murder than, "What must become of the world, if such practices prevail?" There are no cases of these things in which one could think that no harm would be done *in them*, that is, in acts of cruelty, assault, and murder, that no harm *at all*,

and not merely no harm *on balance*, would be done in doing these acts. Which is not to say that generalization arguments have no relevance to 'rules of benevolence'. For example, while each breach of the commandment, 'thou shalt not kill,' harms, it can be only the general breach that would rob "people [of] security of their lives" (Harrison 1981, p. 62).

'Generalization questions' contrast with 'role-reversal questions' or 'Golden Rule Arguments' such as, "How would you like it if someone twisted *your* arm just for the fun of it?" Role-reversal questions are especially, though not exclusively, relevant to rules of benevolence, such as rules against assaults, whereas generalization-questions, when clearly distinguished from role-reversal ones, are relevant especially, even if not exclusively, to 'rules of justice'.

I have, for a mark of 'rules of justice', contrasted cases in which generalization questions have natural relevance, with cases in which, though not without relevance they can seem odd. Let me add that there are of course other cases in which they are of no relevance at all, and for which, outside of academic discussions, they are never contemplated. There are no rules of any kind against being a hairdresser, or going for drives on Sunday. And there are no occasions outside of academic discussions for the questions, 'What if everyone were a hairdresser?' and 'What if everyone went for drives on Sundays?'

3.2 'Extraordinary cases'

3.2.1 Hume tells us that,

"[no one scruples] in extraordinary cases, to violate all regard to the private property of individuals, and sacrifice to public interest a distinction, which had been established for the sake of that interest...The safety of the people is the supreme law...public safety and interest" (E195 [196]).

He cautions to *possibly* similar effect in another place that

"nothing less than the most extreme necessity ...can justify individuals in a breach of promise, or an invasion of the properties of others" (E202 [206]).²³

²³There is extreme necessity of another kind. Hume writes that "where the society is ready to perish from extreme necessity...every man may now provide for himself by all the means, which prudence can dictate, or humanity permit" (E187 [186]). These are supposed to be circumstances of extreme necessity for a society in which "the strict rules of justice are

(continued...)

What exactly are we to make of these 'extraordinary-case' and 'extreme-necessity' (E203 [206]) *caveats*?

3.2.2 *My proposal for extraordinary cases.* Clearly a mere advantage to public interest – that is, that one could do more good than harm by violating a rule of the type of a rule of justice – cannot make a case an 'extraordinary' one. For if it did, justice would never make a *difference*, and constrain benevolence: if a mere advantage were sufficient to make a case 'extraordinary,' it would be right to abide by 'rules of justice', only when they agreed with the supreme law of benevolence. It is for an 'extraordinary' case necessary that one be able to do *much* more good than harm by violating a 'rule of justice.' The question is *how much* more good than harm is needed to make a case, under a 'rule of justice,' an 'extraordinary case' in which the rule can be violated? Hume could, and I think should, say that what is required is that one would do so much more good than harm by violating the rule, that, even if everyone who was in a position to violate 'rules of justice' to public advantage as great were to do so, so *few* would be involved that the rules would remain widely enough observed to serve in their manner the public interest. What is required is that far from suffering, the public interest would in fact enjoy a net gain from these relatively few useful violations taken together. Hume's idea could have been, and I think should have been, that in an '*extraordinary*' case under a 'rule of justice', the straight utilitarian argument against conforming to the rule is *great*, that it is *so great* that there are *few* cases under 'rules of justice' in which straight utilitarian arguments against conforming to them are as great, *so few* that, given the level of conformity to 'rules of justice' in his society, these few cases could, in the interest of the public, be treated as *exceptions* to these rules.

An 'extraordinary' case is by dictionary definition a *kind of unusual* case. On the present reading, an 'extraordinary case' is, for Hume, a case in which the straight utilitarian argument for violating a member of an established and socially useful system of 'rules of justice' is *unusually great, unusually* enough, and *great* enough, that observance of 'rules of justice' in this system *without* observance in these cases, would be better for the public, than the observance of 'rules of justice' that will take place. It can be seen that whether a case is an 'extraordinary case' in a society depends not only on the strength of a straight utilitarian argument for violation, and on how often there would be arguments of at least that strength for violations of 'rules of justice', but also on what has been

²³(...continued)
suspended" (*ibid.*). Hume has in mind *here* situations in which the practice of these rules has *collapsed*, and in which there *are no* established socially useful rules to be obeyed. On (E203 [206]) he writes of extreme necessities that individuals can be under that would have them violating established and socially useful rules.

termed *the surplus of justice in the society* (Section 12.2.4 of the previous chapter), that is, on how much more conformity to 'rules of justice' there is than is needed for the good that general conformity to 'rules of justice' can do. According to the theory I am attributing to Hume, in precarious societies, in which there is barely enough justice (that is, this virtue and the observance of rules it entails) to sustain them, persons are under more stringent 'duties of justice', than are members of more secure societies in which there is much 'justice to spare'. 'Extraordinary cases' in precarious societies need to be extraordinary *indeed*.

3.2.3 Harrison writes:

"It [is] not...that there ought to be no exception to keeping rules of justice. If the heavens were to fall as a result of keeping of a rule of justice then, doubtless, one ought to break it. But, in this case, if everybody were to keep the rule in similarly dire circumstances, the consequences would, like the consequences of just one person's keeping it, also be disastrous. Perhaps a rule of justice should be kept [in a case] if and only if the consequences of everybody's keeping it [as dire a case] are better than the consequences of everybody's breaking it [in these cases]...." (Harrison 1981, p. 71)

I am proposing something similar to Harrison's speculation of a good Humean theory of 'dire-circumstances exceptions' to 'rules of justice'. *One* difference is that I propose a single 'dire-circumstance exception' that would be one and the same for every 'rule of justice', whereas he proposes exceptions that would be specific to individual 'rules of justice'.²⁴ An idea more along Harrison's line for a 'dire-circumstance test' for a 'rule of justice' R, is whether the consequences of keeping *this 'rule of justice'* would be better if *it*, R, were kept in all and only circumstances not so dire. The question he could have one ask is, Would consequences be better *if 'rule of justice' R were violated* whenever circumstances are this dire? I propose the 'dire-circumstance test' for a 'rule of justice' R, that the consequences of keeping *all 'rules of justice'* would be better if *they* were kept in all and only circumstances not so dire. The question I would have asked is, Would consequences be better *if all 'rules of justice' were violated* whenever circumstances are this dire?

²⁴Another difference is that his idea does not make relevant how *common* are 'circumstances as dire'.

A rule by rule test could license exceptions to sundry 'rules of justice', such that were they *all* taken, the consequences would be worse, possibly much worse, than if none were taken. The logical point (explained in Section 6 of Chapter VIII below) is the same as for the possibility that individual exceptions to a single 'rule of justice' could each be good in terms of what would be its consequence, though the consequences were they all taken would be bad. For a sense of the point one may consider a case in which each of two people who can walk across a lawn will not do so, and would not do so, even if the other were to do so. Suppose, as could be, that each, were he to walk across the lawn, would do good on balance, and perhaps good even to the lawn by stimulating its growth; suppose that each person's crossing would do as much good as would the other person's. Even so it could be that if everyone similarly placed were to walk across the lawn the result would be not good, but bad. For I have arranged that there is no one else, and that they are 'similarly placed', so that if everyone similarly placed were to walk across the lawn, then *both* of them would do so, and that could be too much traffic for the lawn to bear.

3.2.4 Caveats *or* caveat? Is there properly one *caveat* implicit in Hume's texts, see Section 3.2.1 above, or two? 'Extraordinary cases' I take to be defined exclusively in terms of the *public* interest. Cases of 'extreme necessity' would, presumably, be defined in terms of the agent's *private* interests? Paradigms for 'extreme necessity' are life-boat cases in which unless someone is eaten all will die! Suppose, to bring a 'rule of justice' into the case, that those in the boat have entered into a compact to do all they can (short of offering themselves up to be eaten!) to keep everyone else alive. There is one person who is failing and will surely die before they are rescued *unless* they come to his aid. If they do that, he will hang on until they are all past surviving, even by cannibalizing him. Their straight utilitarian arguments for breaking that compact are great, but *not necessarily* great enough to be unusual enough for their case to be an 'extraordinary case' as defined in 3.2.2. For it is possible for several acts each to have very bad (very good) consequences, though taken together they have very good (very bad) consequences: Chapter VIII argues for and explains this possibility.²⁵

Hume says that "nothing less than the most extreme necessity can **justify** individuals in a breach of promise, or an invasion of the properties of others" (E202 [206]), *implying* that the most extreme necessity *does*

²⁵*Suppose* otherwise. Suppose all violations of rules of justice in cases of extreme necessity 'taken together' would have good consequences. It does, I think, not strictly follow, simply as a matter of logic, that conformities to rules of justice in all the remaining cases would, 'taken together', have better consequences than would conformities to rules of justice in all cases including these of extreme necessity would, 'taken together', have.

justify 'acts of injustice' in every case. I think he goes too far, given the basis in *public* utility that he endorses for the virtue of justice. What he can accommodate under the head of extreme necessity are two things that in different ways stop short of that principle of unlimited license when under extreme necessity. They are, (i), that extreme necessity justifies when, as in the real world it can be expected always to do, it makes a case an extraordinary case, and, (ii), extreme necessity in *every* case, if it does not justify, then at least *excuses* violations of 'rules of justice', and renders their agents unblameable. The thought for (ii) is that it is, after all, only human to 'buckle' under pressure of the most extreme necessity.²⁶

3.2.5 *Socrates's problem in the Crito*. What could Hume have said to Socrates, if he had been present during Socrates' conversation with Crito? How might Hume have viewed the *form* of Socrates's problem? Pretending that Hume had arrived in time to speak to him, he might have said to Socrates:

'Your problem is, as you have stated, whether it would be right to disobey the state and escape from this place. Your problem is whether it would be right in this case to violate your *duty of allegiance* (E202 [205]), your duty to obey your state's every command. The question therefore comes down to this. Is your case an *extraordinary* case under this 'rule of justice', and so a proper exception to it? Suppose for the moment that you have a substantial straight-utilitarian reason for escaping. (If you do not have a straight-utilitarian reason for escaping, substantial or otherwise, then you have no *moral* problem at all.) The question then is whether it is so substantial and strong that reasons that are that strong for disobedience are sufficiently unusual as to recommend

²⁶Hume is I think contemplating a 'defense' when charged with 'violations of rules of justice' analogous to the 'defense of necessity' of somewhat uncertain scope that can be available against charges of violations of criminal laws. Here, for interest, is a paragraph drawn for the *Encyclopedia of Britannica Online* about that defense, with bold emphasis added.

"The use of force may also be **excused** if the defendant reasonably believed himself to be acting under necessity. The doctrine of necessity in Anglo-American law relates to situations in which a person, confronted by the overwhelming pressure of natural forces, must make a choice between evils and engages in conduct that would otherwise be considered criminal. In the oft-cited case of *U.S. v. Holmes*, in 1842, a longboat containing passengers and members of the crew of a sunken American vessel was cast adrift in the stormy sea. To prevent the boat from being swamped, members of the crew threw some of the passengers overboard. In the trial of one of the crew members, the court recognized that such circumstances of necessity may constitute a defense to a charge of criminal homicide, provided that those sacrificed be fairly selected, as by lot. Because this had not been done, a conviction for manslaughter was returned. The leading English case, *Regina v. Dudley and Stephens*, 14 Q.B.D. 273 (1884), appears to reject the necessity defense in homicide cases. In German or French courts, however, the defendants would probably have been acquitted."

themselves by their utility for all (including yourself) as exceptions to rules such as this one, whose foundation is the public interest in their general observance.

'Don't ask, What would happen were state-orders to count for nought? Ask instead, What would happen were they to count for nought *in cases such as this one*, and similarly for other demands of justice, in cases in which reasons for not meeting them are as *strong* as they are in this one?'

The theory I am attributing to Hume has him say, as did Socrates, that justice and obedience come first before all things. It accords to justice and obedience priority in determinations of what is right. But, unlike Socrates' position in the *Crito*, the theory I am attributing to Hume does not accord to justice and obedience *absolute* priority, or say that one's actions must be just and obedient, no matter *what* the consequences.

Would a correct application of this theory of right action have made a *difference* in Socrates' case? Would it have found that to escape was the right thing to do? Socrates might have said, No, for he seems to have felt that he did not have even *one* good consequential reasons for escaping, let alone consequential reasons that were on balance extraordinarily good. Without demurrer he allows the Laws to say to him, "if you depart...[you injure] those you should injure least -- yourself, your friends, your country and us" (*Crito* 54c). But Socrates could have been wrong about these possible consequences, and, more interestingly, he might have seen his errors, and changed his mind, if he had thought that consequences of escaping for himself, his friends, his family, and Athens at large, if *very* good on balance, could make a difference to the *rightness* of escaping. For then, when determining whether or not it would be right for him to escape, he could not take as a consequence of his escape, that he would need to avoid well-governed cities, or that he would be ashamed, and not able to speak of "virtue and justice as man's most precious possessions" (*Crito* 53, Grube translation), or that he would be a bad influence on his friends and children, *all* of which reflections would be predicated on issue of rightness having been already settled against escaping..

Socrates might have changed his mind, if he had been convinced that to determine what was right he needed to ask the questions that this theory I am attributing to Hume would have had him ask. It is possible that

deliberate application of Hume's theory of right action, in which the priority of rules is *moderated* as indicated by considerations of consequences, would have made a life-saving difference in Socrates' case.²⁷

3.2.6 Hume indicates that in his view justice consists of commitments to honour rules of justice of one's time and place, except in extraordinary circumstance. He does not explain this important reservation. I have proposed an explanation for the extraordinary-case exception, along with a revision of the apparently different extreme-necessity exception.

What are the 'right' explanations, Hume would say, are questions of fact, difficult questions of fact the evidence for which must be found in how – during actual deliberations and discussions of sophisticated 'moralists in the street' – matters of the 'extraordinary', and of 'extreme necessity', are handled. Perhaps all that we can be really confident of is that *some* such reservations are a part of justice as ordinarily understood and realized, and that in this regard some great philosophers – Socrates and Kant – have been simply *mistaken*. Who, other than such theorists, have ever said, and really *meant*, that it is justice – that it is telling the truth, honouring one's agreements, obedience to authorities of one's state, and so on – no matter *what* the consequences? Socrates, it is I think fair to say, should have known, and in his conversation with Crito said, better than that. It is *not*, in common views of persons of good sense, that simple.

"But Socrates and Kant do not claim merely to describe and apply sophisticated common sense, so in their uncompromising rigour they are not mistaken about it. They mean to be saying something general not about prevalent opinions about justice, but about the real thing, when they say its demands, the demands of justice and duty, hold, absolutely no matter what the consequences."

Fair enough. So it is not as if they and Hume disagree about some simple or not so simple fact. They and he are talking about different things. Of course, Hume would wonder how they suppose themselves to have learned this surprising thing about 'real justice', and for that matter what this 'real justice' of which they pretend to speak might be. He would think it is a *chimera* made of senses of duty 'talked up in private dialogues *hyperbolically*', that is, worked up in philosophic reveries far beyond the cool measures of common sense with which, like everyone else, Socrates and Kant must have begun their thinking about justice and duty. And then, worked up, 'out they came'

²⁷Socrates's case is discussed in Chapter II of (Sobel 2005b).

projected as absolute and fully objective, and *then projected*, again and again, so that there was in the end, in their final visions of absolute duties, less than they thought and felt.

4. *Texts that can suggest modifications*

4.1 (E219n [228n], bold emphasis added)

"**Tendencies** of actions...not their **real accidental consequences**, are alone regarded in our moral determinations."

Hume is thinking here, as in most places, about spectator-appraisals of qualities of mind as virtues or vices, and of actions as virtuous or vicious. But I assume that he would say that determinations to do, and to favour actions (never mind motives), even by ideally well-informed agents would proceed similarly in terms of tendencies and likelihoods of various consequences. I have taken it to be part of the interpretive assumption of Section 1.2 that leads to the theory attributed to Hume in Section 2.1, that not even an ideally well-informed agent would *know*, would have certain and true beliefs about, what would be the consequences of sundry possible actions – that even a fully informed virtuous agent would, of necessity, act, and base his favour for actions, and his advice to others, on probabilities, albeit ideally well-informed probabilities.

Right actions in the theory I attribute to Hume are identified with what an ideally well-informed perfectly moral agent would favour and advise, and so for reasons just stated this theory runs at a point in terms of likelihoods of possible consequences. But the theory could be changed in this connection to another, consistent perhaps, with the letter, even if not the spirit of Hume's books. The theory could be changed to one projected on the slightly different interpretive assumption (and considerably less plausible *interpretive* assumption, I think) that identified right actions with what a perfectly moral agent *with prescience*, with certain and true beliefs, about what would be the consequences of possible actions, would favour and advise. [*Cf.* Smith's analytic line that we have reasons to do what fully rational agents would want us to do (Smith 1994, p. 151), wherein the 'summary notion' of 'full rationality' includes having no false beliefs, and having all relevant true beliefs (p. 156).]

The objectifying revision of the previous paragraph in the interpretive assumption leads to a correspondingly different principle for right actions: a principle suited to that revision comes from that of Section 2.1 by the following boldly emphasized changes.

An action A is right in a case C

if and only if:

EITHER

(for one thing) there is at least one rule R that applies in case C such that, (i), R is a 'contributing member' of an 'established' and 'socially useful' system of 'rules of justice' for the time and place of C, and, (ii), C is not an 'extraordinary case' in relation to R,

(for a second thing) action A conforms to every rule that satisfies (i) and (ii), and

(for a third thing), of actions that conform to every rule that satisfies (i) and (ii), action A would, **in view of what would be its consequences**, be at least as attractive from the standpoint of general happiness and utility as would be any other action that is open to the agent that satisfies (i) and (ii);

OR

(for one thing) there is no rule R such that R satisfies both of conditions (i) and (ii), and

(for a second thing) action A would, **in view of what would be its consequences**, be at least as attractive from the standpoint of general happiness and utility as would be any other action that is open to the agent.

A theoretical advantage of this objectified theory of right actions is that it makes room for theories of subjectively right actions of two sorts. One would be, for actions that are right given the agent's actual probabilities, and the other for actions subjectively right given ideally well-informed probabilities. The theory of right action of Section 2.1 is identical with what would be this second subjective-elaboration of this objective principle.

4.2 *(E217n [225n])*

"In proportion to the station which a man possesses, according to the relations in which he is placed, we always expect from him a greater or less degree of good...When the interests of one country interfere with those of another, we estimate the merits of a statesman by the good or ill, which results to his own country."

Similarly, one supposes, for ordinary citizens, with due allowance to their more limited opportunities for doing good for their countries. And so one can wonder whether tests of utility in Hume's theory of right should be, not of general utility for humankind, but rather of local utility for agents's communities. But the theory as it stands is on this point suited to his books. He says that the *end* of morality is quite unrestricted public utility and "the true interests of

mankind" (E183 [180]). His view is that, even so, the best *means* is for persons to take as their ends more definite, and in details more familiar, objects.

"[N]ature has implanted in every one a superior affection to his own country....Not to mention that, while every man consults the good of his own community...**the general interest of mankind** is better promoted, than by any loose indeterminate views of the good of a species [our species]...."

(E217n [225n], bold emphasis added)

Partiality, Hume is saying, is not only natural but *useful* – it is useful *for getting to, or at least approximating, right actions*, when, subject to applicable constraints (if any) of rules, one is trying to do what is best. It is a good thing for ordinary persons of limited information to be somewhat partial to family, friends, and neighbours.²⁸ However, ideal moral agents – fully informed agents possessed of every moral virtue in proper measure – would give scrupulously impartial moral advice, and it is their advice that according to my reading of Hume's intentions determines what is morally right and wrong. When considerations of utility are relevant – for example, when 'rules of justice' are not decisive) – they would advise for general and total, rather than merely local and partial, utility.

"It is wisely ordained by nature that private connexions should commonly prevail over universal views and considerations; otherwise our affections and actions would be dissipated and lost, for want of a proper limited object. Thus a small benefit done to ourselves or our near friends excites more lively sentiments... than a great benefit done to a distant commonwealth. But still we know here...to correct these inequalities by reflection, and retain a general standard of vice and virtue founded chiefly on **general usefulness**." (E220n [229n])

Ideal moral advice is founded entirely on completely, all mankind, general usefulness, though in a complicated manner that pays due attention to the particular species of utility of justice. (*Cf.*, Baron 1988.)

²⁸Hume does *not*, as Aristotle might, suggest that to compensate for what I recognize to be my partiality to my community, I should give *more* weight than seems to me appropriate to the interests of members of other communities. Consider the *Nicomachean Ethics* II(9), 1109a19-1109b28, titled by M. Ostwald, "How to attain the mean." Discussed in Section 4.3 of Chapter VII in (Sobel 2005c).

4.3 (E265 [289])

"When a man at any time, deliberates concerning his own conduct – as, whether he had better, in a particular emergence, assist a brother or a benefactor – he must consider these several relations, with all the circumstances and situations of the persons, in order to determine the superior duty and obligation."

There is nothing in this passage that suggests the details of deliberation appropriate to the mixed actual-rule/act utilitarian theory of right that I have attributed to Hume, or even its utilitarianism. But then there is nothing in this passage that is inconsistent with that theory, which can be taken to detail *how*, in what *order*, and to what *ends*, such a man – a man who for whatever reason is *deliberating* concerning his own conduct – is to *consider* the relations, circumstances, and situations of persons in his case.

5. *Difficulties for the theory I attribute to Hume*

The utilitarian theory I attribute to Hume is troubled by possibilities of *conflicts of duty*. These may call for modifications to the theory. They might, if noticed by Hume, lead to retractions of, or qualifications to, things he says. The theory has another problem with *bad moralities* which problem may call for 'adulterations' of its utilitarianism. And this theory makes trouble for radical utilitarian spirits – for very *good* people, one may think – who are wont to serve the general happiness every time without reserve. It makes trouble for them of a kind made by every 'indirect utilitarianism'.

5.1 *Conflicts of duties*. Established socially useful rules of the type of rules of justice are not complicated products of genius. They are relatively simple practices that come to be established somehow naturally to enable cooperation, coordination, and commerce to everyone's advantage. And so it is not surprising that they can conflict in cases, and that they sometimes do conflict in cases, that are not, in the sense explained, 'extraordinary'. When they do conflict, the theory of right proposed for Hume in Section 2.1 says that *nothing* is right!

5.1.1 Suppose, for example, that I have promised to meet you for lunch, and find myself, with no time to lose if I am to keep that promise, on the wrong side of someone else's field. There is, assume, no way I could have predicted my predicament. The area is new to me. I have done nothing wrong, but there I am. Either I break my promise, or trespass, and so, one way or the other, I violate a rule of the type of a rule of justice. Furthermore, it is clear that the

sky wouldn't fall if I were late. You would be inconvenienced, and for a time somewhat unhappy, but the case is far from being extraordinary or particularly unusual in the strength of my straight-utilitarian reasons for trespassing, rather than keeping you waiting.

This case may be felt to embarrass the theory I have attributed to Hume. Since there is a rule that satisfies conditions (i) and (ii) of the first disjunct of the principle, only that disjunct applies. But by the 'second thing' of that disjunct an action is right only if it conforms to the *rules* that satisfy conditions (i) and (ii) – only if it conforms to *every one of them*. And so in the present case this principle says that no action is right. That can seem unsatisfactory. It is as if this principle had been framed in blissful innocence of the real world and the possibilities of conflicts of duties. I note that there is in fact evidence that Hume sometimes overlooks just such possibilities. For he says:

"Nothing less than the most extreme necessity, it is confessed, can justify individuals in **a breach of promise, or an invasion of the properties of others.**" (E203 [206])

This statement has the clear, albeit easily over-looked consequence, that in a non-extraordinary case such as I have described, in which there is no extreme necessity, and obligations to keep promises and duties to respect property are at odds, there is nothing that the agent can be justified in doing, or excused for doing. The clear implication is that in such a case, his choice in the case is between actions *each* of which violates a rule of justice proper, and is wrong and, as said, inexcusable..

5.1.2 Several reactions are possible to the anomalous behaviour of the theory in conflict-of-duty cases. *One reaction* would be to say that in such cases nothing *is* morally right. This reaction says that morality, given that part of it is determined by established useful rules of the type of rules of justice, is such that it is possible for a person, through no fault of his, to find himself in situations from which there is no moral escape – situations in which there is nothing that he can do that would be morally right, situations in which whatever he does will be morally wrong.

Alternative possible reactions would consist of changes to the theory of Section 2.1, so that there was in every case something that, according to the theory, would be right. A simple change to this effect would take us to a theory that made an action right in a case like the one described, if it satisfied *at least one* of the rules covering the case, even if it did not satisfy *every one*. That could not recommend itself to Hume.

Other changes that might appeal more to Hume, would – in a case in which no action satisfies every rule that meets conditions (i) and (ii), and several actions satisfy at least one such rule – have the test of utility applied directly to these several actions to decide amongst them. That could be *how* “a man [who] deliberates...whether he had better...assist a brother or a benefactor...must consider...in order to determine the superior duty and obligation” (E265 [289]), and, whether or not ‘rules of justice’ are said to clash in this case, that could be how Hume would have a man decide, when they do, which makes the superior duty. Hume holds that public safety and interest is the supreme law, and that we are to go to it in extraordinary cases (E195 [196]). Perhaps he would say that we are also to defer to it, in the manner I have suggested, to resolve conflicts of rules of the type of rules of justice, and that we are to defer to it in these cases whether or not there is enough at stake to make the case an extraordinary one, or one of extreme necessity. If so, we need to say for him, which of two kinds of applications of the supreme law are to decide. The procedure could be, in a conflict-of-duties-of-justice case, to test for utility alternative rule-following actions in the case, or to test for utility alternative practices of like rule-following actions, in all similar conflict-cases.

5.1.3 Of the three reactions to notice of the possibility of conflict-of-duty cases, I *suspect* that the first would be Hume's choice. He could say that, like it or not (!), that is of the nature of morality. We can get ourselves into moral corners of completely ordinary circumstances from which there is no way out without violating an established useful ‘rule of justice’. And when we do, *there is no right way out*, and *every way out is wrong*. Perhaps Hume did *not* overlook the possibility of conflicts between duties to keep promises and to respect others property. Perhaps he meant what his words on (E203 [206]) imply, namely, that when these duties conflict, and neither is distinguished by opposition to it of an ‘extraordinary case’, then no way is morally justified, the agent ‘can do no right’, *the agent must do wrong*. Hume could add that when an agent finds himself in a moral *cul-de-sac* through no fault of his, then he is of course not at fault, and is not to be blamed or thought less of for his bad luck and the wrong he does. [For discussions variously congenial to this position see contributions by Ruth Marcus and Norman Dahl in (Mason 1996).]

5.2 'Bad moralities', for example, of institutionalized slavery. Rules of exploitation, and dispositions to honour them, that, though they would be to the advantage of many, grind down the oppressed who resent them, and actively oppose them, so that all in all these rules are a *drag* on the general happiness, that would be better without them – *such* rules cannot be objects of informed and impartial sympathetic approval. Such rules, on informed and impartial views of their affective consequences, including “effects of their resentment” (E190 [190]), will not be found better than no rules in their areas. They will not be 'contributing members' of established socially useful systems of 'rules of justice.' Transgressions of them will not be wrong according to the actual-rule clause of the mixed utilitarian principle for right and wrong actions that I attribute to Hume.

Probably the rules for slavery in mid-nineteenth century states of the union shortly before the Civil War were of this character, so that at that time it would have been better for the general happiness that there should be no rules for slavery, that there should be no lawful practice or condition of slavery.

"But what if, at some time and place, an institution of slavery in which slaves are made miserable, is up and running? What if there is an institution of slavery that is not resented and actively resisted by slaves, but accepted as defining their 'natural lot', an institution that is not under attack by social reformers, and an institution that is not opposed by any sentiments of humanity of the subjugating class, perhaps because they 'see' their slaves as children? What if there is an institution of slavery in circumstances such as *that*, an institution of benign slavery acquiesced in so that 'lords' do not “feel the effects of their [slaves's] resentment” (E190 [190], *cf.* the boxed note in Section 3.1.1 above) since they do not resent it, though it exploits and makes them miserable, that is serving the *general* happiness better than it would be served by in place 'arrangements of justice and their rules' *minus* this institution and its rules? Isn't this possible not only logically but really, psychologically and sociologically? Isn't this, though far-fetched, 'just possible'."

I think so.

"And wouldn't that make honouring its rules, and for example turning in run-away slaves (of whom by assumption there would be few), *right*, save in extraordinary circumstances, according to the theory that has been attributed to Hume?"

Yes.²⁹

What should we make of that? Not, I think, that *the attribution* of that theory to Hume is after all mistaken. The evidence for it still seems strong. The problem is with *the attributed theory*. The problem is that we are apt to suppose that honouring the rules of any institution of slavery that makes slaves miserable would be *wrong*, even if the institution is serving the *general* happiness. Many of us would say (though some of us, uneasily) that honouring the rules of an institution of slavery is wrong, *no matter where or when* and *no matter what*, and so even if slaves in it are *happy* and accepting of it, even if they are in favour of it.

I see no good way of dealing with this problem on Hume's behalf by changing the theory, without compromising the 'purity' of his utilitarian intentions evidenced by many passages: I think that Hume must say to those of us who suppose that full participation in a social/economic system that includes an exploitative institution of slavery, even when the system is stable, and serving the general happiness thanks in part to the presence in it of this institution, would be wrong, that *we* are wrong. He might add that full participation in a system is consistent with speech against parts of it to change it, and, say that when speech can change rules for the better, he is for it. (But what if there are laws against 'seditious speech against the system', laws of the type of rules of justice against it?) Hume might quote, in agreement, that one is 'to persuade or obey' (*Crito* 51a-c). He might say, 'Participate, do no wrong, even while, if there are prospects for some good out of it, you lobby by all lawful means for better social/political/economic arrangements.' (But again there is the possibility of laws against that, in order to bolster 'the system', a system that by hypothesis is presently in everyone's interest, partly because no one minds it.) Hume just might go further and encourage in circumstances of 'extraordinary opportunity' unlawful means to promote arrangements that will, after a period of turmoil and social distutility, serve better the general interest of humanity, provided that the transitional disutility was made up by the gains secured. Though in this advocacy he could not, consistent with the theory I find in his texts, say that employing of such means would be morally justified. It would be his humanity speaking out, that would have us 'go outside of not only law but morality'.

In all this I speak not of what Hume did, but only of what he might, say about situations of bad moralities. This is because in his books and essays on morals he actually says little about the important topics of the processes,

²⁹Thanks to Seka Jankovic for pressing these questions.

the social/psychological dynamical principles, of living and *changing* systems of social rules and institutions, and about moral principles concerned with *changing and reforming* institutions and regimes. Though he does have a fair bit to say about circumstances for lawful *resistance and rebellion* in "Of the Measures of Allegiance" in the *Treatise* (T110-13 [549-53]).

5.3 *Humanity's challenge to morality*

"[Virtue] declares that her sole purpose is to make...all mankind...happy...." (E258 [279])

"The safety of the people ['public safety and interest'] is the supreme law: All other particular laws are subordinate to it, and dependent on it...." (E195 [196])

Why then not pursue this purpose, and go directly to the supreme law in every case? Why isn't the rule in every case simply to promote utility as best one can? If, as Hume says, the greatest possible promotion of the general happiness is the supreme end of morality, why doesn't a moral and virtuous person get on with it, as best he can, without handicaps of rules? Why doesn't the fully informed paragon of moral virtue advise that we get on with it, promotion of the general happiness, in every case as best we can? Why isn't the way to right action always simply to apply the test of greatest happiness for humanity directly to actions one can do, then to identify actions that score best on this test? "I mean, that's what it's all about right? Making all of mankind happy."

Hume's answer could be this – words in single quotation-marks are his, all others are mine that I imagine him speaking:

"*In my moral philosophy I DESCRIBE, I do not advocate.* The first issue for me is what a truly moral and virtuous person would be like. From this we can say things about what she would, if fully informed, do, how she would deliberate, and what she would favour and advise that we do.

"One of the first things to say regarding what a truly moral and virtuous person would be like, is that she would be *honest and just.* *Who will deny this?* But in her honesty and justice, to begin a list of her 'artificial virtues' (all of which share the species of utility that distinguishes justice from benevolence), she would refer her actions to established socially useful rules, as has been explained, and advise that all do likewise. Dispositions of scrupulousness to defer, and to favour deference by all, to such established rules are what honesty and justice and the other 'artificial virtues' *are* as qualities of minds.

"The virtuous person would 'find' that established rules of the type of justice generated obligations and duties for her, and for others, *moral* duties and obligations that can run counter in particular cases to what would best serve his private interests, and, not infrequently, counter also to what would best serve the public interest. *And she would not be fazed by this*. The moral virtue of this person is by hypothesis in tact, and she has every moral virtue including the *main* one, *justice*, with the species of utility of which she would *familiar*. Perhaps she takes comfort and strength for this virtue of hers, and her disposition integral to it to abide by its rules, from the intelligence that the public interest is best served by general observance of these, even though particular observances are not always, and need not even often be, in the public interest.

"It is a difficult truth about morality that single acts of justice are 'frequently contrary to *public interest*....[as well as] private interest'; it is a hard truth about them that single acts of 'justice may be pernicious in every respect' (T65-6 [497-8]). It is *the hard truth about morality* that acts that are pernicious in every respect can be *morally correct and required*, for, as said, acts of justice in entirely ordinary circumstances can be like that. Acts of justice are *frequently* like that. *With* its peculiar species of utility, justice is a *virtue, made so* I say by the general usefulness for humanity of its being wide-spread and strong. And *acts* of justice are generally morally correct and required. *Who would deny this?!*

"I prefaced these remarks with a disclaimer. I said that *in my moral philosophy* I describe, and do not advocate. But I freely confess that I can discover in what I describe only what I *personally approve*, and, *in my life*, out of love of humankind, *advocate*."³⁰

5.4 *Robin Hood's Problem*. Suppose that Hume is correct about the *de facto* structure of morality, and that morality as we know it *does* accord a measured and qualified priority to the demands of certain socially useful rules. Suppose that it does sometimes require conformity to rules in circumstances in which acts of conformity considered singly are hurtful and positively disuseful for all affected. If this is the truth about morality, and in particular about its 'rules of

³⁰Boris Pasternak writes of his character, Colonel Strelnikov, that: "In order to do good to others he would have needed, besides the principles which filled his mind, an unprincipled heart, the kind of heart that knows of no general cases, but only of particular ones, and has the greatness of small actions. Filled with the loftiest aspirations from his childhood, he had looked upon the world as a vast arena where everyone competed for perfection, keeping scrupulously to the rules." (*Doctor Zhivago*, pp. 226-7). The trick, solved for us by our sympathetic and rational natures which have established the virtues of benevolence and justice, is 'to have it both ways'. The problem for a virtuous and reflective person is to sustain and live with his Janian character which especially in troubled times can be frequently conflicted.

justice', then even the most public spirited person, indeed *especially* the most public spirited person, can wonder *why he should be moral, and in particular, why he should be just*. It is not only the "sensible [selfish] knave" (E260 [282]) who has this problem. It is also a problem for a *benevolent humanitarian extremist* who, even if just and so bothered by his plans to steal from the rich to give to the poor, would be a *sensible generous knave*. It is a problem for *Knights of Humanity* who would be *absolutely unscrupulous in the public interest*. *Robin Hood's Problem* is of a piece with problems left for kind-hearted loan officers, and do-gooder lawyers in the previous chapter. Each situation, according to Hume's view of "that species of utility that attends [justice]" (E262 [285]), is a way in which Glaucon's Problem (*Republic* 357a - 367b) can be real for a lover of humanity who has a passion for the general happiness and sees the rules of justice of his day getting in the way of his pursuit of that end. The business of justice and its rules can be not only *complicated* in practice, as said in Section 2.2 above, but, we are now reminded, trying and *difficult*, especially for 'the man of every virtue' including in *generous* measure *benevolence*.

6. Looking Ahead

I believe that Hume must say that it can be *reasonable* for really extreme partisans of humanity to be unjust 'in their hearts', to be *ready* to lie and cheat and steal and in general to flaunt 'rules of justice' whenever general happiness can be furthered thereby, and to be willing to pay the price for this preparedness. By this I do not mean the price exacted by society. I mean the price they will pay for their being willing to engage in benevolently motivated unscrupulousness, the price they will pay in having 'bad characters with themselves' and with the loss of that "peaceful reflection on one's own conduct" that is reserved for the just and morally scrupulous (E261 [284]). It is a price that Hume argues that these extremists will pay even if they have Rings of Gyges. Explaining this interested obligation that almost everyone has to virtue, makes, however, another story a 'footnote' to which is that *Knights of Humanity, though not 'conflicted' as virtuous and reflective people sometimes are (see the previous note), can be expected to be not as happy as these people*. Before getting to it I take a closer look in Chapter VII at that ever troubling 'species of utility of justice.' Challenges reiterated to the possibilities, logical and real, of this species of utility are the subjects of Chapters VIII and IX. Then, better prepared, Chapter X tells Hume's impressive management of the mother of all questions in moral philosophy, now given a Humean twist, Why be moral, why in particular – given the troubling species of utility of this virtue – *be just?*

Appendix A. Contractarian Considerations

Now come comments on 'contractarian moments' in Hume's texts to do with justice, moments that are not reflected in the theories of right action and of ideal moral deliberation that I attribute to him, which hew to the lines that:

"**The good of mankind** is the **only object** of all laws and regulations by which justice is directed, and property determined.... [T]he rules which we follow...are such as can best be contrived to serve farther **the interests of society**." (E192 [192], text rearranged and bold emphasis added.)

A1. *Mutuality of interest, and when 'rules of justice' take place and direct reasonable persons*

At least rules of justice proper – rules of property, truth-telling, and promise-keeping – have that species of utility that consists in practices of them by many, though often not their individual observances, having public utility, and indeed also, for each person, private utility. Furthermore, there will probably be a certain *mutuality* of interest in practices of these rules. In Hume's view one can expect there to be 'something in' at least these 'rules of justice' that are to take precedence in the deliberations of individuals that would recommend them to each member of a group of persons, as terms of agreements with *all* others in this group – one expect something for everyone in these rules *in order that they should take place*, on the assumption that they can take place for a community only if they would be *acceptable* to its members as terms of an agreement. In a community of unequals wherein "a species of creatures intermingled with men...were possessed of such inferior strength, both of body and mind, that they were incapable of all resistance" (E190), there would not be a common interest in the establishment of rules for mutual regards and forbearance by all towards all, and so – assuming accurate discernment of interests, and a good measure of rationality in resolutions and behaviour – the process of 'convention' described in the *Treatise* (T59-60) would not operate to produce such rules of unrestricted scope, but would operate amongst only the strong to establish *rules of limited scope* for mutual regards and forbearance by them towards their equals, and possibly, operating amongst only the weak – those inferior creatures and 'natural slaves', other rules and licenses, as well as virtues, which could include complacency in their 'second-class citizenship' and subordination to the human superiors among whom they lived, and to whom they were subordinate. The inferior creatures that Hume imagines in order to make this point about justice, would not, his language and discussion indicates, be *human* beings, or at any rate would be very inferior specimens of human beings (E190-1 [190-1]).

A2. Utilitarian and contractarian considerations compared

There are in Hume's thoughts about actions not only utilitarian considerations throughout of public utility, but also, specifically on the justice side, and for actions that would reflect this virtue, certain *contractarian aspects of mutuality*. All 'rules of justice' established in a community will not only have the peculiar species of utility of which he so often writes (see Section 1.2.2 above, and the next chapter), but, since they are in place, they can be expected to define an arrangement that even in its full scope tends to be of 'mutual interest'. Hume is *at pains* to distinguish questions "*concerning the manner in which the rules of justice are established,*" and questions "*concerning reasons which determine us to attribute to the observance or neglect of these rules a moral beauty and deformity*" (T55 [484], italics are Hume's). The rules of justice, in their peculiar manner, promote "the true interests of mankind" (E183 [180]), possess "public utility," and have the "beneficial consequences ...[that] are the *sole* foundation of [the] merit" of justice, and the sole grounds for our moral approbation of it (E185 [183]). In contrast, the recently indicated difficult condition of mutual interest in established 'rules of justice', in Hume's view contributes *not at all* to their *merit* and our approval of them, but is instead important – assuming accurate views and a good measure of rationality – merely *to their taking place*, and connectedly, once they are in place, to their directing the conduct of *reasonable* individuals (*cf.*, E191 [191]). The theories that I attribute to Hume of right action and of ideal moral deliberation to act, and on reflection to advise, say, by significant omissions, that *only* the public utility of 'rules of justice' – and not also that there is a mutuality of interest in them – figures as a condition of right action, and is looked for during ideal *moral* deliberations and assessments. I suggest that according to Hume, considerations of common interest, and of mutual interest, enter only into deliberations regarding whether it is *reasonable* to take direction in a case from 'rules of justice', and that in his implicit view it need not be reasonable to regard rules in this manner and take direction from them, even when they are *morally decisive* and it would be *wrong* not to take direction from them in a case.³¹

³¹What then of the following, found in Hume's argument to show that there is not an exceptionless duty to society that a suicide would transgress?

"All our obligations to do good to society seem to imply something reciprocal. I receive the benefits of society, and therefore ought to promote its interests...." (Hume 1993, p. 320: "Of Suicide:" submitted to a printer in 1755, to be withdrawn along with "Of Immortality." Both were with Hume's permission published posthumously.)*
Hume writes here of *all* our obligations *to do good to society*. He has in mind obligations we may have to *contribute* to society in 'our work'. *These* cease when benefits of society to me cease. He does not say that 'duties of justice', for example, to respect the property of others, to be truthful, and to be faithful to promises, and so on cease. It is relevant that in his view honouring
(continued...)

Suppose I were to think that some people amongst us are not being exploited by the system, and that, because they are strong and resourceful, *they* would be better off were our system of property and commerce to *collapse*, and for us to be left ‘propertyless’. It would be *wrong*, according to the descriptive/explanatory theory of morality that I find in Hume’s texts, for me to steal from the rich, to give to these poor, and *wrong* for *them* to help themselves, though it *might* not be *unreasonable* for either of us to oppose the system in these ways. If I am right in my reading of Hume here, there comes a difficult question: Is his theory *true* on this point of ‘morality as found’? I think it is.

A3. *A troublesome text*

There is, I confess, a passage in the *Treatise* that troubles my reading of Hume. In that passage, which runs in Section IX of the *Treatise* from "Few persons can carry on this train of reasoning" (T112 [552]) to the end of that section, Hume says that:

"The moral obligation [of allegiance to a government] is founded on the natural [interested obligation], and therefore must cease where *that* ceases; especially where...[we] foresee very many occasions wherein the natural obligation may cease....if interest first produce obedience to govern-

³¹(...continued)

these duties is frequently *not* ‘good to society’: honouring these duties is, he allows himself to say, frequently pernicious in every respectg.

*Further to “On Suicide”: Most of this essay argues that there is not a duty to God not to take one’s own life. G. R. McLean argues that despite “a widespread view among philosophers that Hume’s attack upon [theistic objections to suicide] succeeds, when in fact...this attack does not succeed at all” (McLean 2001, p. 100a.). I think he exaggerates, though it is curious that Hume does not consider ‘head on’ the popular argument that we are not ours to dispose of because we our God’s creatures. Suppose that I make a pot, and that you come along and smash it. That is wrong because it was not yours to destroy. Similarly, the argument goes, it is not for us, but for God, to dispose of us; for He made us, and we are his. Cf.: “men being all the workmanship of one omnipotent, and infinitely wise make...they are his property” (Locke, *The Second Treatise Government*, Chapter 2.) “God is our owner, we are his property” (Kant 1963, p. 154.)

This popular argument is troubled by the reflection that the theology presupposed would have God creating all living things, mosquitoes, pine trees, bacteria, viruses, all of that. In that context the argument implies, *absurdly*, that we encroach on God’s province when we take antibiotics. Adapting words of Hume’s, he could to bolster this ‘*reductio ad absurdum*’, say that “[i]n order to destroy the evidence of [it], we must show a reason why this particular case [of our disposing of *ourselves*] is excepted. It is because human life is of such great importance, that it is a presumption for human prudence to dipose of it? But the life of man is of no greater importance to the universe [and would be of no more importance to its Creator, Hume implies] than that of an oyster” (p. 319). Of course, in Hume’s view, human beings *are* of greater *value* and *more important* – in the hierarchy of ‘human morality’ – than the animals we hunt and eat, in the case of oysters, eat alive. “But,” he implies, “Divine Morality and ‘morality enshrined in reality at large’ would presumably not incorporate the oh so human biases of ours.”]

ment, the obligation to obedience must cease whenever the interest ceases in any great degree and in a considerable number of instances." (T113 [552-3])

Hume takes for granted here that when private individual interest ceases *often*, and *in great degree*, then necessarily (and not merely probably) so does "the interest of society" (T112 [552]) and "the public interest" (T113 [553]). But *these* are terms he sometimes employs for *foundations of the moral merit* of adherence to rules, and not, as here, for "our original motive to obedience [to them]" (T113 [553]). Perhaps Hume connects the two senses and thinks that when the private individual interests of members of the public cease, often and in great degree, then necessarily so does the public interest in the sense of those "beneficial consequences...[that] are the *sole* foundations of [the] merit" of obedience (E185 [183]). *Perhaps*, though that opinion is not consistent with his view that obedience to a government has *the species of utility of justice*. A practice or system of actions has that species of utility only if, though it, this practice or system, is in the public, and even in every person's private interest, *every* individual act of obedience in it can be contrary to both public and private interest, and presumably contrary in great degree – the degree of contrariness, even if relevant to its probability, should not be relevant to the possibility. But whatever should be the final word on this troublesome passage in the *Treatise* – the book that, in his Advertisement to the second volume of the posthumous edition of 1777 of *Essays and Treatises*, Hume tells us *to ignore* – and whether or not, as I think, he would say that this difficult passage contains "negligences...in reasoning" (E173 [2]); it does not imply that a person has no moral obligation to obey, whenever he has no interested obligation to obey. It is said only that "[moral] obligation to obedience must cease whenever the interest ceases in any **great degree** and in a **considerable number** of instances" (T113 [553], bold emphasis added). *Cf.*:

"The observance of justice, though useful among [nations], is not guarded by so strong a necessity as among individuals; and the *moral obligation* holds proportion with the *usefulness*.... All politicians will allow...that reasons of state may in particular emergencies dispense with rules of justice, and invalidate any treaty or alliance where the strict observance of it would be prejudicial in a considerable degree to either of the contracting parties. But nothing less than most extreme necessity, it is confessed, can justify individuals in a breach of promise, or an invasion of the properties of others." (E203 [206])

A4. *On the purity of Hume's utilitarianism*

That Hume's implicit theory of right action is *holus bolus* contractarian is certainly wrong, and no one thinks it, since there is plainly in Hume's view much more than justice to being moral. There is also benevolence with which convention, and conditions of reasonable agreements, have absolutely nothing to do. That his implicit theory of right action is not even in part, and on the justice side, contractarian – that it is purely utilitarian – is indicated by many texts including prominently the first paragraph of the section "Of Justice":

"That Justice is useful to society, and consequently that *part* of its merit, at least, must arise from that consideration, it would be superfluous to prove. That public utility is the *sole* origin of justice, and that reflections on the beneficial consequences of this virtue are the *sole* foundations of its merit; this proposition ...will better deserve of our examination and enquiry." (E185 [183]).

Hume is saying in this passage that public utility, or usefulness to and beneficial consequences for society, is the sole 'origin' of the virtuousness of that quality of mind that is justice, the sole 'foundation' for its moral merit – it is that, its public utility that 'gives it the power of virtue'. Hume's point here is not new in the *Second Enquiry*: *cf.*, "justice is a moral virtue, **merely** because [that is, *exactly* for the reason that] it has [a] **tendency to the good of mankind**" (T133 [577]), and "[j]ustice is...approved of [and meritorious] **for no other reason** than because it has **a tendency to the public good**" (T167 [618], bold emphasis added). Considerations of 'immediate agreeableness' either for the just, or for those with whom they interact, have nothing to do with it. (That is one difference, though not for present purposes an important difference, Hume finds between justice and its subdivisions, and benevolence.³²) I go from the idea that public utility is the sole foundation of the virtue (virtuousness) of justice, through the peculiar *species* of that utility of this virtue, to the conclusion in the principle of right action that I have assembled for Hume, that it is only the public utility of established rules of justice that is relevant to our moral obligations under them, and that it is, contrary to David Gauthier, not also required that 'right-making' rules should be "mutually advantageous" (Gauthier 1990, p. 60), that they should have "*mutual expected utility*...[with] each person reasonably expect[ing] that rule or practice to be useful to himself" (*op. cit.*, p. 57).

³²The merit of benevolence, arising from its utility, and its tendency to promote the good of mankind, has been already explained [in Section II: see (E183 [181])], and is, no doubt, the source of a *considerable* part of that esteem...But it must be also be allowed, that...[sentiments of benevolence are] delightful in themselves [and by sympathy] communicated to spectators" (E241 [257]).

We must, Hume could remind, keep separate, questions concerning the actually origins of beneficial arrangements, and questions concerning the foundations of the morality of qualities of mind supportive of these arrangements, as well as questions concerning our obligations given such arrangements. Thus: "[A] regard to public interest...is not our first and original **motive** for the observation of the rules of justice...." (T64 [495-6], bold emphasis added.).

"The *natural* obligation to justice, *viz.*, has been fully explain'd; but as to the *moral* obligation, or the sentiment of right and wrong.... (T67 [498].)

"[I]nterest...the *first* obligation to the performance of promises. Afterwards a sentiment of **morals** concurs with interest, and becomes a new obligation upon mankind. This sentiment of morality of promises arises from the same principles as that in the abstinence from the property of others. *Public interest, education, and the artifices of politicians*, have the same effect in both cases." (T88 [523], bold emphasis added.)

This new sentiment of morals *concurs* with the first interested obligation, and arises after it, but it neither 'arises *from*' it, nor consists in part of it. It arises from reflection on public interest, aided by education and artifices of politicians. This "*moral* obligation, or...sentiment of right and wrong" (T67 [498]) is, to propose Kantian ornamentation, what the sentiment of self-directed moral approbation becomes in cases in which there are *not* plain and palpable concurring interested obligations, and *especially* in cases in which there are plain and palpable *opposing* interested obligations. Such cases can, according to Hume, be frequent under rules that have that species of utility that attends justice. There is discussion of 'sentiments of right and wrong' and 'senses of duty' in Hume, of their origins and uses, in Section 4 of Chapter IX below.

A5. '*Public utility*' vs '*mutual expected utility*'

A5.1 Gauthier, in his very useful essay, "David Hume, Contractarian" (Gauthier 1990), with which I do not entirely disagree, is concerned to separate, as Hume would have him do, questions of origins of practices and qualities of

mind, from questions of their morality,³³ so that he may attribute to Hume the view that ‘mutual expected utility’ is required in order for ‘practices’ of property, government, promises, and the like to *obtain* (p. 52), but that ‘general usefulness’ alone, without the additional factor that arrangements “may be expected to be useful to each person....enter[s] into the moral approbation accorded justice” (p. 58, I have rearranged Gauthier’s words). Gauthier allows that this ‘bifurcated line’ on justice in the *Second Enquiry* can seem to be challenged by the second sentence of Section III, “Of Justice,” in the *Second Enquiry*.

“Hume’s account of justice begins [*sic*] with the claim [*sic*] ‘that public utility is the *sole* origin of justice, and that reflection on the beneficial consequences of this virtue are the *sole* foundation of its merit’.” (E185 [183].)

Accepting the challenge, Gauthier writes:

“In reply to this reading of Hume’s claim, I shall insist that public utility is to be understood as *mutual expected utility*....I shall also insist that beneficial consequences are not to be equated with public utility, so that the foundation of merit of justice is to be distinguished from the origin of justice.” (P. 58.) “[Hume’s] initial claim should...[be] read, with words added [thus]: “That public utility [i.e., mutually expected advantage] is the *sole* origin of justice, and that reflections on the generally beneficial consequences of this virtue are the *sole* foundation of its merit [i.e., moral approbation].” (P. 58.)

I am surprised by Gauthier’s response for two reasons.

First, there was available to him a much *easier* response. He could have said that Hume is not in this claim distinguishing, indeed, contrasting, ‘the *origin of justice*’ from ‘the *foundation of its merit*’, but that, as can be

³³Gauthier, however, runs together questions of the actual origin of ‘justice’ in the sense of ‘just social arrangements’ with questions of what is *essential* to just social arrangements *and what makes them just*. While ‘general usefulness’ of an arrangement is sufficient for its ‘receiving moral approval’, he has Hume saying:

they are *just* only if “beneficial consequences extend to each person”; “this additional factor is essential to *justice*”; “it is not the beneficial consequences themselves, but the expectation of benefit by each person, that is just-making” (Gauthier 1990, p. 58).

I have not found Hume saying *anything* about the *justice* of social arrangements. He theorizes about how ‘systems of justice’ come to be, and are maintained, and of course about the moral merit of them and the quality of mind that serves their maintenance. But he does *not*, as far as I can see, have a theory of the justice of social arrangements, a theory of what makes them just. I suppose he has an *idea* of possible ‘systems of justice’, of in particular of possible systems of property, good, and bad, capable of being not only established, but maintained, by freely interacting rational agents (!!), and of systems not so capable. But he does not, I think, have, as Gauthier *seems* to say he has, ‘a theory of justice’ of a kind that can be utilitarian, contractarian, Rawlsian, or whatever.

gathered from his emphasizing not these differing words, but the same word, ‘sole’, he was, with his theory of virtue as a power of qualities of mind to elicit approval on reflection, elaborating in the second conjunct, the substance of the first conjunct. Gauthier could have said, as I do say, that the second sentence in “Of Justice,” is to be read this way:

“That public utility is the *sole* origin [of the virtue (the virtuousness)] of justice, and that reflections on the beneficial consequences are the *sole* foundation of its merit; this proposition...will...deserve our examination and enquiry.” (E185 [183].)

Second, there is no evidence in Hume’s text for the proposition that he at least *sometimes* meant by ‘public utility’ what Gauthier means by ‘mutual expected utility’, and there is decisive evidence regarding the text that provoked Gauthier’s contrary claim that Hume did not use ‘public utility’ in that sense there. This evidence emerges when the *context* of the words Gauthier interprets is recalled. After a brief survey of Hume’s uses of ‘public utility’ in the next section, I ‘context’ the words that troubled Gauthier in Section 4.5.3 below to pose the question of his interpretation of them. Had Hume wished to intend something like ‘mutual expected utility’ on some occasions when he uses ‘public utility’, he would have had ready better words than these, for example, ‘common interest and advantage’, or ‘mutual interest and advantage’ (T86 [520]), or ‘common interest and *utility*’ which occurs in the *Second Enquiry*:

“War has its laws as well as peace; and even that sportive kind of war, carried on among wrestlers, boxers, cudgel-players, gladiators, is regulated by fixed principles. Common interest and utility beget infallibly a standard of right and wrong among the parties concerned.” (E206 [211].)

A5.2 Cole Mitchell (see note in Section 1.1 of the previous chapter) has found fifteen occurrences of ‘utility’ in Book III of the *Treatise*, one occurrence of ‘public utility’ (T94 [531]), and one of ‘publick utility’ (T71n [504n]). That use of ‘public utility’ does not come with a ‘local gloss’. ‘Motives of public interest’ are said to be “derived from publick utility.”

There is one occurrence of ‘utility’ in “A Dialogue”. It stands in an occurrence of ‘public utility’ that is without a local gloss: (E[334]).

Mitchell has found fifty eight occurrences of ‘utility’ in the *Second Enquiry*, fourteen of which stand in occurrences of ‘public utility’, and one in ‘public interest and utility’. Regarding the thirteen other than the one that

bothered Gauthier, twelve are without obvious ‘local glosses’:(E198 [201]), (E200 [203])³⁴, (E203n [207n]), (E204 [208]), (E208 [214]), (E277 [305]: two occurrences), (E278 [306]), (E279 [308]), (E280 [309]), (E281n [309n]), (E281n [310n])!! ‘Public utility’ is glossed on (E183 [180]) with the indication that it is “decided...by ascertaining, **on any side**, the true **interests of mankind**.” On (E201 [293]) there occurs ‘public interest and utility’: the ‘interest’ is here said to be extreme – “The necessity of justice to the support of society is the sole foundation of that virtue” (*ibid.*). While not bearing explicitly on his use of ‘public utility’, it may be relevant to it that Hume writes, regarding the great relevance of ‘utility’ to “the praise, which we bestow on the social virtues,” that it is not supposed “that a greater eulogy can be given to any man, than to display his **usefulness to the public**” (E207 [212], bold emphasis added). Again, Hume writes of an unlikely person who has absolutely no concern for “the public good of a community or the private utilities of others” (E224 [235]): if the public good of a community was ‘the mutual expected utility of its members’, it would be constituted of their private expected utilities. Also, courage is said to have utility “to the public and to the person” (E238 [254]).

A5.3 Gauthier takes on what he sees as the *prima facie* challenge to his interpretation of Hume on justice in these words that make up the first half of the second sentence in his section “Of Justice”:

“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” (E185 [183]).

Gauthier’s interpretation of this sentence includes taking ‘public utility’ in it to mean what Gauthier terms ‘mutual expected utility’. We have seen that Gauthier can find no evidence for this reading in *other* uses of ‘public utility’ in Hume’s major work on morals. His case must at least start with this very use, and what we can see it does, or may, mean *here*. Is *this* use of ‘public utility’, unlike most uses that Hume makes of this term, ‘locally glossed’? Let us begin by completing the sentence in which it stands. It may provide a clue.

“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; this proposition, **being more curious and important**, will better deserve our examination and enquiry.”

³⁴Here the public utility of *justice* is raised to the level of, my term, ‘public necessity’, regarding breaches of rules of justice, “we...ask, *What must become of the world, if such practises prevail? How could society subsist under such disorders?*” (E200 [203].)

No help yet, though it is clear where to look next, for we want to know, More curious and important than what? The answer is , The proposition expressed by the *first* sentence in “Of Justice”:

“That Justice is useful to society, and consequently that *part* of its merit, at least must arise from **that consideration**, it would be superfluous undertaking to prove.”

We are ‘getting warmer’. *That consideration is being useful to society*, a consideration that, in the last two paragraphs in “Of Benevolence,” the just previous Section II, is said to be part at least of what pleases us about benevolence and social virtues of its ilk.

“Upon the whole...it seems undeniable....*that a part*, at least, of [the sentiment of benevolence’s] merit arises from its **tendency to promote the interests of our species, and bestow happiness on human society**. The social virtues are never regarded without their beneficial tendencies....**The happiness of mankind**, the order of society, the harmony of families, the mutual support of friends, are always considered as the result of their gentle dominion over the breasts of men.

“How considerable a *part* of their merit we ought to ascribe to their **utility**, will better appear from future disquisitions [footnote: “Sect. III and IV”] ; as well as the reason, why this circumstance has such a command over our esteem and approbation [footnote: “Sect. V”].” (E183-4 [181-2]).

So ‘that consideration’ of the first sentence in “Of Justice” is the utility referenced in the last of the previous section, which is the tendency to promote the interests of our species, and bestow happiness on human society. *That*, Gauthier would be the first to agree, is not ‘mutual expected utility’. Which brings us to the question of his interpretation: Is it the *least* plausible that stylist and word-smith Hume should have juxtaposed as he does the proposition that (here comes a paraphrase), ‘at least *part* of the merit of justice arises from **that utility**’, the claim of the first conjunct of a ‘more interesting’ proposition, which is ‘that **public utility** is the *sole* origin of justice’, though he did *not* mean by ‘public utility’ the kind of utility he had so recently referenced (with the words ‘that consideration’ of course, not the words ‘that utility’ of my paraphrase)? You be the judge.

Appendix B. 'Of the Influencing Motives of Right and Wrong'

B1. *Thoughts of right and wrong cannot alone move to action:*

the question is, "What is still missing?"

The issue is not whether *judgments of right and wrong* carry with them motivation. Hume can, consistent with his strictures against motives that would be provided by reason alone, say that judgments of right and wrong, when sincere, carry with them motivation. Such judgments for Hume are always at least in part approvals and disapprovals, and so when 'sincere', or at least when *heartfelt*, they *do* carry with them motivation. And such judgments are generally, if not always, also in part ascriptions of *fictitious* qualities that would be objectively and authoritatively prescriptive. Hume can plausibly maintain that *these* ascriptions entail motivation so that, as a believer in these qualities has said, "[t]o feel I ought to pay my bills," to *appreciate, recognize, apprehend* this obligation, "is to be moved towards paying them" (Prichard 1970, p. 90). H. A. Prichard uses 'feel' (pp. 88, 90, 94), 'appreciate' (pp. 90, 91, 92, 96), 'recognize' (pp. 91, 92), and 'apprehend' (pp. 91, 96) interchangeably in relation to obligations.

The issue is whether *thoughts of right and wrong*, thoughts of *possible facts* of right and wrong, carry with them motivation, when these thoughts are not inspired by projective errors, and are of propositions that have nothing to do with fictitious qualities. Plain facts of rights and wrongs, I have assumed, are for Hume certain facts about ideal moral advice spelled out in Section 1.1 above. Suppose, therefore, for an issue peculiarly relevant to Hume as I read him, that I think that, as a matter of fact, if an ideally well-informed and perfectly moral person were to offer me sincere moral advice, he would advise me to keep my promise to you. Believing that, might I even so not be moved, not be at all moved, to keep my promise to you? Might this be so, even after reason has done all its work not only *for* this thought, but *on* it, to tap its potency for me? Yes, that is possible, and Hume must agree, for he says (famously) "that reason alone can never be a motive to any action of the will" (T23 [413]).³⁵

³⁵By 'action of the will' Hume does *not* here mean precisely and literally actions of the *will*. He is not speaking precisely of choices, decisions, resolutions, preferences and the like. He is sure that "reason alone can never...give rise to **volition**," but he says not only that, but that it can never "**produce any action** or give rise to volition" (T24 [414], bold emphasis added). His subject, indicated in that most famous line by the words 'action of the will', is precisely *motivated* action in contrast with reflexive, instinctive, and perhaps also habitual action. His subject is motivated *action in the world* such as handing over a sum of money to pay a debt, or buying a plum (T24 [414]). These, incidentally, are Hume's *only* examples, and even they are merely implicit, of the 'actions of the will' with which he is mainly concerned in the section "Of the Influencing Motives of the Will."

Hume 'says' – if I may be allowed once more to extrapolate from his words – that it would be possible, for a person who was not the least lacking in capacities of reason to reflect, compare, calculate, and so on, after he had done everything possible with rational capacities on a thought, *any* thought, to plumb its potency, to be quite unmoved and unimpressed by it. And even if there are possible thoughts that make exceptions to this general principle of the impotence of reason alone, for example, thoughts of ideal advice *tout court*; Hume is surely right about the thought in hand concerning ideal *moral* advice.

Arguing, decisively I think, even for the *general* principle, without exceptions, is that "[t]hrough spiritual or physical tiredness, through accidie,..., through general apathy, through despair,...through a feeling of uselessness or futility, and so on" [(Stocker 1979, p. 744), quoted in (Smith 1994, p. 120)] a person can be less and less, to the point of being *not at all*, motivated to do *anything*, and this *regardless of what and how much he thinks about his options*. But the status of the general principle need not be decided for present purposes, for which it is sufficient that, as said, Hume is clearly right about the thought in hand concerning ideal *moral* advice.³⁶

B2. *Might Sympathy Be the Answer?*

"But if you are not moved by this thought of ideal moral advice, *though you are perfectly rational, are fully informed, and have deployed reason to the fullest on your information on this advice to tap its potency for you*, there must be something about you that sets you apart from the common run of humanity."

³⁶The general principle, as I understand it, is, in other fewer words, that it is possible, without being in any way 'practically irrational,' to be unmoved by a thought, any thought. The principle is that no thought passes what has been termed the 'internalism requirement', the requirement that it motivate insofar as the thinker is practically rational (*cf.*, Korsgaard 1986, pp. 11, and 14-5). Witness, as I have said, Stocker's depressed. Smith, who opposes this general principle, writes in effect against my witness for it, that the "various sorts of 'depression' as Stocker calls them" are "practical irrationalities of various kinds" (Smith 1994, p. 120). Later he elaborates to say that the depressed, in their relative, or at the limit total, lack of desire, need not themselves be irrational, but that their 'positive indifferences' are irrational "to the extent that they wouldn't be had [were they not depressed]" (p. 155). The possibility of Stocker's depressed does not, Smith implicitly argues, show that it is possible, without 'practical rationality,' not to be moved by any thought, for these unfortunates, though not necessarily 'practically irrational,' always harbour 'practically irrational indifferences'.

Smith argues, I think 'by misnomer' or redefinition. The depressed, we can say, using words in ordinary ways, and facing up to what we may wish were not true, are *not* necessarily irrational, nor is their state of depression, or many indifferences necessarily irrational. The depressed can be simply *depressed* – not confused, not wildly and strangely mistaken about important things, but simply, and of course unfortunately, depressed. They may, but then again they may not, be able to say why they are depressed. Existential lethargy, to focus on that, is not irrational in itself, and need not be so in its 'grounds,' if any. It is only (only?!) tragic. A person can be, without mistakes or confusion, depressed and weary of life. It happens. Nature's gift, for which we can only be grateful, is that it does not happen more often, that it so often passes of itself when it does happen, and that strategies are available to many of us for hurrying through dark nights.

True, I think, and Hume can agree, which brings the question, “What sets me apart, if I am not moved by thoughts of ideal moral advice, what is it that I am missing?” Something motivational and arational, but what? It is tempting to think that what, in the presently elaborated case, I must be missing is, in Hume's view, *sympathy*, and that this deficiency not merely sets me apart from the rest of humanity, but from *all* of humanity – that this deficiency, indeed, makes me less than human.³⁷ For texts to this temptation we have:

"that which renders morality an active principle...it is probable [in this early passage, Hume is happy to leave 'the controversy' be]...[is] some internal sense or feeling, which nature has made universal in the whole species." (E177 [173])

"If we consider the principles of the human make, such as appear to daily experience and observation, we must...conclude it impossible for such a creature as man to be totally indifferent to the well or ill-being of his fellow-creatures...." (E220 [230])

“[H]e must be more or less than man, who kindles not in the common blaze [as does every man by a particular exercise of sympathy].”³⁸ (E254 [275])

The temptation is to say that on Hume's view, if I think that ideal moral advice would be to do an act, am perfectly rational, have deployed reason to the fullest on this thought to tap its potency for me, and am yet not the least motivated to do it, then *I cannot be 'perfectly human'* – the temptation is to think that, in Hume's view, I must then be somewhat 'denatured' and, because devoid of sympathy, deficient in humanity, 'less than a man.'

"But," it may be countered, "sympathy even combined with perfect reason and information, is not the answer. For even then I could be *amoral* and quite devoid of every moral virtue. I could even then be at least *somewhat* amoral and *deficient* in moral virtue. I could even then be deficient in, and indeed quite devoid of, *justice*. And if lacking in moral virtue, if, to go to the extreme, I am

³⁷The answer to what sets me apart, if I am not moved by thoughts of right and wrong, cannot be that I lack *senses of duty*, for right and wrong actions need not be made so by peremptory or prohibitive ‘rules of justice’, and it is only for such actions that a person of perfect moral virtue sometimes needs and has for motives ‘thoughts of duty’. Such thoughts have nothing to do with a *kind* person’s avoidance, as he steps to his seat in a theatre, of the feet of those who have taken their seats before him.

³⁸more than...a man? Would the reference be to angels out of sympathy with living humans, or to a god out of sympathy with them? There is evidence, in Part XII of Hume's *Dialogues Concerning Natural Religion*, that, for want of a better hypothesis to explain appearances of order and design in the world, Hume may have believed in such an indifferent god.

completely lacking in it, then I may well not be on an occasion the least interested in ideal moral advice for me in my circumstances. If I am lacking in moral virtue, then I may *not* think less of myself, I may *not* suffer a 'loss of character with myself' – *cf.*, (E261 [283]), when I fail to do what I realize that a fully informed paragon of morality would have me do. If lacking in moral virtue, I may lack not only every *moral* motivation to do right and avoid wrong 'when no one will know', but every *interested* obligation. For I may not be interested in 'going through the motions of virtue', in order to get it."

The premise of this counter is denied in Chapter X below. According to the argument of that chapter, if I am fully informed, and fully reasonable, *and sympathetic*, then, though I need not be moral, though I need not be kind and honest and the rest, I cannot be *amoral*. According to that argument, if I am fully informed and reasonable and sympathetic, then I will at least be somewhat *for* morality in myself as well as in others, and *will* pay a price in loss of self-esteem to the extent that I see myself as somewhat deficient in its virtues, and whenever I ignore what I accept as sound moral advice, which I will for this reason be *somewhat* disinclined to do, which is to say, somewhat moved not to do.

Sympathy *is*, for Hume, the answer to the question, "What am I *missing*, if, though informed and reasonable, I am not at all moved by thoughts of ideal moral advice?" It is the answer to that question according to the argument of Chapter X which, while addressed to the question, "Why be of a perfectly moral character?" has implications, as indicated, for the question, "What of value am I missing, if, though I am not perfectly moral, I do not think less of myself for that?"³⁹

³⁹I am grateful to Richmond Campbell, Susan Dimock, Ann Levey, and especially Willa for helpful criticisms and suggestions for matters of this chapter..