Hume's moral philosophy

Hume's position in ethics, which is based on his empiricist theory of the mind, is best known for asserting four theses: (1) Reason alone cannot be a motive to the will, but rather is the "slave of the passions" (see Section 3) (2) Moral distinctions are not derived from reason (see Section 4). (3) Moral distinctions are derived from the moral sentiments: feelings of approval (esteem, praise) and disapproval (blame) felt by spectators who contemplate a character trait or action (see Section 7). (4) While some virtues and vices are natural (see Section 13), others, including justice, are artificial (see Section 9). There is heated debate about what Hume intends by each of these theses and how he argues for them. He articulates and defends them within the broader context of his metaethics and his ethic of virtue and vice.

Hume's main ethical writings are Book 3 of his Treatise of Human Nature, “Of Morals” (which builds on Book 2, “Of the Passions”), his Enquiry concerning the Principles of Morals, and some of his Essays. In part the moral Enquiry simply recasts central ideas from the moral part of the Treatise in a more accessible style; but there are important differences. The ethical positions and arguments of the Treatise are set out below, noting where the moral Enquiry agrees; differences between the Enquiry and the Treatise are discussed afterwards.

1. Issues from Hume’s Predecessors
2. The Passions and the Will
3. The Influencing Motives of the Will
4. Ethical Anti-rationalism
5. Is and Ought
6. The Nature of Moral Judgment
7. Sympathy, and the Nature and Origin of the Moral Sentiments
8. The Common Point of View
9. Artificial and Natural Virtues
10. Honesty with Respect to Property
   - 10.1 The Circle
   - 10.2 The Origin of Material Honesty
   - 10.3 The Motive of Honest Actions
11. Fidelity to Promises
12. Allegiance to Government
13. The Natural Virtues
14. Differences between the Treatise and the Moral Enquiry
Bibliography
Academic Tools
Other Internet Resources
Related Entries

1. Issues from Hume’s Predecessors

Hume inherits from his predecessors several controversies about ethics and political philosophy. One is a question of moral epistemology: how do human beings become aware of, or acquire knowledge or belief about, moral good and evil, right and wrong, duty and obligation? Ethical theorists and theologians of the day held,
variously, that moral good and evil are discovered: (a) by reason in some of its uses (Hobbes, Locke, Clarke), (b) by
divine revelation (Filmer), (c) by conscience or reflection on one’s (other) impulses (Butler), or (d) by a moral sense: an
emotional responsiveness manifesting itself in approval or disapproval (Shaftesbury, Hutcheson). Hume sides with the
moral sense theorists: we gain awareness of moral good and evil by experiencing the pleasure of approval and the
uneasiness of disapproval when we contemplate a character trait or action from an imaginatively sensitive and
unbiased point of view. Hume maintains against the rationalists that, although reason is needed to discover the facts of
any concrete situation and the general social impact of a trait of character or a practice over time, reason alone is
insufficient to yield a judgment that something is virtuous or vicious. In the last analysis, the facts as known must trigger
a response by sentiment or “taste.”

A related but more metaphysical controversy would be stated thus today: what is the source or foundation of moral
norms? In Hume’s view this is the question what is the ground of moral obligation (as distinct from what is the faculty for
acquiring moral knowledge or belief). Moral rationalists of the period such as Clarke (and in some moods, Hobbes and
Locke) argue that moral standards or principles are requirements of reason — that is, that the very rationality of right
actions is the ground of our obligation to perform them. Divine voluntarists of the seventeenth and eighteenth centuries
such as Samuel Pufendorf claim that moral obligation or requirement, if not every sort of moral standard, is the product
of God’s will. The moral sense theorists (Shaftesbury and Hutcheson) and Butler see all requirements to pursue
goodness and avoid evil as consequent upon human nature, which is so structured that a particular feature of our
consciousness (whether moral sense or conscience) evaluates the rest. Hume sides with the moral sense theorists on
this question: it is because we are the kinds of creatures we are, with the dispositions we have for pain and pleasure,
the kinds of familial and friendly interdependence that make up our life together, and our approvals and disapprovals of
these, that we are bound by moral requirements at all.

Closely connected with the issue of the foundations of moral norms is the question whether moral requirements are
natural or conventional. Hobbes and Mandeville see them as conventional, and Shaftesbury, Hutcheson, Locke, and
others see them as natural. Hume mocks Mandeville’s contention that the very concepts of vice and virtue are foisted
on us by scheming politicians trying to manage us more easily. If there were nothing in our experience and no
sentiments in our minds to produce the concept of virtue, Hume says, no lavish praise of heroes could generate it. So to
a degree moral requirements have a natural origin. Nonetheless, Hume thinks natural impulses of humanity and
dispositions to approve cannot entirely account for our virtue of justice; a correct analysis of that virtue reveals that
mankind, an “inventive species,” has cooperatively constructed rules of property and promise. Thus he takes an
intermediate position: some virtues are natural, and some are the products of convention.

Linked with these meta-ethical controversies is the dilemma of understanding the ethical life either as the “ancestors” do,
in terms of virtues and vices of character, or as the “moderns” do, primarily in terms of principles of duty or natural law.
While even so law-oriented a thinker as Hobbes has a good deal to say about virtue, the ethical writers of the
seventeenth and eighteenth centuries predominantly favor a rule- or law-governed understanding of morals, giving
priority to laws of nature or principles of duty. The chief exception here is the moral sense school, which advocates an
analysis of the moral life more like that of the Greek and Hellenistic thinkers, in terms of settled traits of character —
although they too find a place for principles in their ethics. Hume explicitly favors an ethic of character along “ancient”
lines. Yet he insists on a role for rules of duty within the domain of what he calls the artificial virtues.

Hume’s predecessors famously took opposing positions on whether human nature was essentially selfish or
benevolent, some arguing that man was so dominated by self-interested motives that for moral requirements to govern
us at all they must serve our interests in some way, and others arguing that uncorrupted human beings naturally care
about the weal and woe of others and here morality gets its hold. Hume roundly criticizes Hobbes for his insistence on
psychological egoism or something close to it, and for his dismal, violent picture of a state of nature. Yet Hume resists
the view of Hutcheson that all moral principles can be reduced to our benevolence, in part because he doubts that
benevolence can sufficiently overcome our perfectly normal acquisitiveness. According to Hume’s observation, we are
both selfish and humane. We possess greed, and also “limited generosity” — dispositions to kindness and liberality
which are more powerfully directed toward kin and friends and less aroused by strangers. While for Hume the condition
of humankind in the absence of organized society is not a war of all against all, neither is it the law-governed and highly
cooperative domain imagined by Locke. It is a hypothetical condition in which we would care for our friends and
cooperate with them, but in which self-interest and preference for friends over strangers would make any wider
cooperation impossible. Hume’s empirically-based thesis that we are fundamentally loving, parochial, and also selfish
creatures underlies his political philosophy.

In the realm of politics, Hume again takes up an intermediate position. He objects both to the doctrine that a subject
must passively obey his government no matter how tyrannical it is and to the Lockean thesis that citizens have a
natural right to revolution whenever their rulers violate their contractual commitments to the people. He famously
criticizes the notion that all political duties arise from an implicit contract that binds later generations who were not party
to the original explicit agreement. Hume maintains that the duty to obey one’s government has an independent origin
that parallels that of promissory obligation: both are invented to enable people to live together successfully. On his view,
human beings can create a society without government, ordered by conventional rules of ownership, transfer of
property by consent, and promise-keeping. We superimpose government on such a pre-civil society when it grows large
and prosperous; only then do we need to use political power to enforce these rules of justice in order to preserve social
cooperation. So the duty of allegiance to government, far from depending on the duty to fulfill promises, provides
needed assurance that promises of all sorts will be kept. The duty to submit to our rulers comes into being because
2. The Passions and the Will

According to Hume's theory of the mind, the passions (what we today would call emotions, feelings, and desires) are impressions rather than ideas (original, vivid and lively perceptions that are not copied from other perceptions). The direct passions, which include desire, aversion, hope, fear, grief, and joy, are those that “arise immediately from good or evil, from pain or pleasure” that we experience or think about in prospect (T 2.1.1.4, T 2.3.9.2); however he also groups with them some instincts of unknown origin, such as the bodily appetites and the desires that good come to those we love and harm to those we hate, which do not proceed from pain and pleasure but produce them (T 2.3.9.7). The indirect passions, primarily pride, humility (shame), love and hatred, are generated in a more complex way, but still one involving either the thought or experience of pain or pleasure. Intentional actions are caused by the direct passions (including the instincts). Of the indirect passions Hume says that pride, humility, love and hatred do not directly cause action; it is not clear whether he thinks this true of all the indirect passions.

Hume is traditionally regarded as a compatibilist about freedom and determinism, because in his discussion in the *Enquiry concerning Human Understanding* he argues that if we understand the doctrines of liberty and necessity properly, all mankind consistently believe both that human actions are the products of causal necessity and that they are free. In the *Treatise*, however, he explicitly repudiates the doctrine of liberty as “absurd... in one sense, and unintelligible in any other” (T 2.3.2.1). The two treatments, however, surprisingly enough, are entirely consistent. Hume construes causal necessity to mean the same as causal connection (or rather, intelligible causal connection), as he himself analyzes this notion in his own theory of causation: either the “constant union and conjunction of like objects,” or that together with “the inference of the mind from the one to the other” (ibid.). In both works he argues that just as we discover necessity (in this sense) to hold between the movements of material bodies, we discover just as much necessity to hold between human motives, character traits, and circumstances of action, on the one hand, and human behavior on the other. He says in the *Treatise* that the liberty of indifference is the negation of necessity in this sense; this is the notion of liberty that he there labels absurd, and identifies with chance or randomness (which can be no real power in nature) both in the *Treatise* and the first (epistemological) *Enquiry*. Human actions are not free in this sense. However, Hume allows in the *Treatise* that they are sometimes free in the sense of ‘liberty’ which is opposed to violence or constraint. This is the sense on which Hume focuses in ECUH: “a power of acting or not acting, according to the determinations of the will;” which everyone has “who is not a prisoner and in chains” (ECUH 8.1.23, Hume’s emphasis). It is this that is entirely compatible with necessity in Hume’s sense. So the positions in the two works are the same, although the polemical emphasis is so different — iconoclastic toward the libertarian view in the *Treatise*, and conciliatory toward “all mankind” in the first *Enquiry*.

Hume argues, as well, that the causal necessity of human actions is not only compatible with moral responsibility but requisite to it. To hold an agent morally responsible for a bad action, it is not enough that the action be morally reprehensible; we must impute the badness of the fleeting act to the enduring agent. Not all harmful or forbidden actions incur blame for the agent; those done by accident, for example, do not. It is only when, and because, the action’s cause is some enduring passion or trait of character in the agent that she is to blame for it.

3. The Influencing Motives of the Will

According to Hume, intentional actions are the immediate product of passions, in particular the direct passions, including the instincts. He does not appear to allow that any other sort of mental state could, on its own, give rise to an intentional action except by producing a passion, though he does not argue for this. The motivating passions, in their turn, are produced in the mind by specific causes, as we see early in the *Treatise* where he first explains the distinction between impressions of sensation and impressions of reflection:

An impression first strikes upon the senses, and makes us perceive heat or cold, thirst or hunger, pleasure or pain, of some kind or other. Of this impression there is a copy taken by the mind, which remains after the impression ceases; and this we call an idea. This idea of pleasure or pain, when it returns upon the soul, produces the new impressions of desire and aversion, hope and fear, which may properly be called impressions of reflection, because derived from it. (T 1.1.2.2)

Thus ideas of pleasure or pain are the causes of these motivating passions. Not just any ideas of pleasure or pain give rise to motivating passions, however, but only ideas of those pleasures or pains we believe exist or will exist (T 1.3.10.3). More generally, the motivating passions of desire and aversion, hope and fear, joy and grief, and a few others are impressions produced by the occurrence in the mind either of a feeling of pleasure or pain, whether physical or psychological, or of a believed idea of pleasure or pain to come (T 2.1.1.4, T 2.3.9.2). These passions, together with the...
instincts (hunger, lust, and so on), are all the motivating passions that Hume discusses.

The will, Hume claims, is an immediate effect of pain or pleasure (T 2.3.1.2) and “exerts itself” when either pleasure or the absence of pain can be attained by any action of the mind or body (T 2.3.9.7). The will, however, is merely that impression we feel when we knowingly give rise to an action (T 2.3.1.2); so while Hume is not explicit (and perhaps not consistent) on this matter, he seems not to regard the will as itself a (separate) cause of action. The causes of action he describes are those he has already identified: the instincts and the other direct passions.

Hume famously sets himself in opposition to most moral philosophers, ancient and modern, who talk of the combat of passion and reason, and who urge human beings to regulate their actions by reason and to grant it dominion over their contrary passions. He claims to prove that “reason alone can never be a motive to any action of the will,” and that reason alone “can never oppose passion in the direction of the will” (T 413). His view is not, of course, that reason plays no role in the generation of action; he grants that reason provides information, in particular about means to our ends, which makes a difference to the direction of the will. His thesis is that reason alone cannot move us to action; the impulse to act itself must come from passion. The doctrine that reason alone is merely the “slave of the passions,” i.e., that reason pursues knowledge of abstract and causal relations solely in order to achieve passions’ goals and provides no impulse of its own, is defended in the Treatise, but not in the second Enquiry, although in the latter he briefly asserts the doctrine without argument. Hume gives three arguments in the Treatise for the motivational “inertia” of reason alone.

The first is a largely empirical argument based on the two rational functions of the understanding. The understanding discovers the abstract relations of ideas by demonstration (a process of comparing ideas and finding congruencies and incongruencies); and it also discovers the causal (and other probabilistic) relations of objects that are revealed in experience. Demonstrative reasoning is never the cause of any action by itself: it deals in ideas rather than realities, and we only find it useful in action when we have some purpose in view and intend to use its discoveries to inform our inferences about (and so enable us to manipulate) causes and effects. Probable or cause-and-effect reasoning does play a role in deciding what to do, but we see that it only functions as an auxiliary, and not on its own. When we anticipate pain or pleasure from some source, we feel aversion or propensity to that object and “are carry’d to avoid or embrace what will give us” the pain or pleasure (T 2.3.3.3). Our aversion or propensity makes us seek the causes of the expected source of pain or pleasure, and we use causal reasoning to discover what they are. Once we do, our impulse naturally extends itself to those causes, and we act to avoid or embrace them. Plainly the impulse to act does not arise from the reasoning but is only directed by it. “Tis from the prospect of pain or pleasure that the aversion or propensity arises...” (ibid.). Probable reasoning is merely the discovering of causal connections, and knowledge that A causes B never concerns us if we are indifferent to A and to B. Thus, neither demonstrative nor probable reasoning alone causes action.

The second argument is a corollary of the first. It takes as a premise the conclusion just reached, that reason alone cannot produce an impulse to act. Given that, can reason prevent action or resist passion in controlling the will? To stop a volition or retard the impulse of an existing passion would require a contrary impulse. If reason alone could give rise to such a contrary impulse, it would have an original influence on the will (a capacity to cause intentional action, when unopposed); which, according to the previous argument, it lacks. Therefore reason alone cannot resist any impulse to act. Therefore, what offers resistance to our passions cannot be reason of itself. Hume later proposes that when we restrain imprudent or immoral impulses, the contrary impulse comes also from passion, but often from a passion so “calm” that we confuse it with reason.

The third or Representation argument is different in kind. Hume offers it initially only to show that a passion cannot be opposed by or be contradictory to “truth and reason”; later (T 3.1.1.9), he repeats and expands it to argue that volitions and actions as well cannot be so. One might suppose he means to give another argument to show that reason alone cannot provide a force to resist passion. Yet the Representation Argument is not empirical, and does not talk of forces or impulses. Passions (and volitions and actions), Hume says, do not refer to other entities; they are “original existence[s],” (T 2.3.3.5), “original facts and realities” (T3.1.1.9), not mental representations of other things. Since Hume here understands representation in terms of copying, he says a passion has no “representative quality, which renders it a copy of any other existence or modification” (T 2.3.3.5). Contradiction to truth and reason, however, consists in “the disagreement of ideas, consider’d as copies, with those objects, which they represent” (ibid.). Therefore, a passion (or volition or action), not having this feature, cannot be opposed by truth and reason. The argument allegedly proves two points: first, that actions cannot be reasonable or unreasonable; second, that “reason cannot immediately prevent or produce any action by contradicting or approving of it” (T3.1.1.10). The point here is not merely the earlier, empirical observation that the rational activity of the understanding does not generate an impulse in the absence of an expectation of pain or pleasure. The main point is that, because passions, volitions, and actions have no content suitable for assessment by reason, reason cannot assess prospective motives or actions as rational or irrational; and therefore reason cannot, by so assessing them, create or obstruct them. By contrast, reason can assess a potential opinion as rational or irrational; and by endorsing the opinion, reason will (that is, we will) adopt it, while by contradicting the opinion, reason will destroy our credence in it. The Representation Argument, then, makes a point a priori about the relevance of the functions of the understanding to the generation of actions. Interpreters disagree about exactly how to parse this argument, whether it is sound, and its importance to Hume’s project.

Hume allows that, speaking imprecisely, we often say a passion is unreasonable because it arises in response to a mistaken judgment or opinion, either that something (a source of pleasure or uneasiness) exists, or that it may be obtained or avoided by a certain means. In just these two cases a passion may be called unreasonable, but strictly
speaking even here it is not the passion but the judgment that is so. Once we correct the mistaken judgment, “our passions yield to our reason without any opposition,” so there is still no combat of passion and reason (T 2.3.3.7). And there is no other instance of passion contrary to reason. Hume famously declares, “’Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. ’Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. ’Tis as little contrary to reason to prefer even my own acknowledg’d lesser good to my greater, and have a more ardent affection for the former than for the latter.” (2.3.3.6)

Interpreters disagree as to whether Hume is an instrumentalist or a skeptic about practical reason. Either way, Hume denies that reason can evaluate the ends people set themselves; only passions can select ends, and reason cannot evaluate passions. Instrumentalists understand the claim that reason is the slave of the passions to allow that reason not only discovers the causally efficacious means to our ends (a task of theoretical causal reasoning) but also requires us to take them. If Hume regards the failure to take the known means to one’s end as contrary to reason, then on Hume’s view reason has a genuinely practical aspect: it can classify some actions as unreasonable. Skeptical interpreters read Hume, instead, as denying that reason imposes any requirements on action, even the requirement to take the known, available means to one’s end. They point to the list of extreme actions that are not contrary to reason (such as preferring one’s own lesser good to one’s greater), and to the Representation Argument, which denies that any passions, volitions, or actions are of such a nature as to be contrary to reason. Hume never says explicitly that failing to take the known means to one’s end is either contrary to reason or not contrary to reason (it is not one of the extreme cases in his list). The classificatory point in the Representation Argument favors the reading of Hume as a skeptic about practical reason; but that argument is absent from the moral Enquiry.

**4. Ethical Anti-rationalism**

Hume claims that moral distinctions are not derived from reason but rather from sentiment. His rejection of ethical rationalism is at least two-fold. Moral rationalists tend to say, first, that moral properties are discovered by reason, and also that what is morally good is in accord with reason (even that goodness consists in reasonableness) and what is morally evil is unreasonable. Hume rejects both these. Some of his arguments are directed to one and some to the other thesis, and in places it is unclear which he means to attack.

In the *Treatise* he argues against the epistemic thesis (that we *discover* good and evil by reasoning) by showing that neither demonstrative nor probable/causal reasoning has vice and virtue as its proper objects. Demonstrative reasoning discovers relations of ideas, and vice and virtue are not identical with any of the four philosophical relations (resemblance, contrariety, degrees in quality, or proportions in quantity and number) whose presence can be demonstrated. Nor could they be identical with any other abstract relation; for such relations can also obtain between items such as trees that are incapable of moral good or evil. Furthermore, were moral vice and virtue discerned by demonstrative reasoning, such reasoning would reveal their inherent power to produce motives in all who discern them; but no causal connections can be discovered *a priori*. Causal reasoning, by contrast, does infer matters of fact pertaining to actions, in particular their causes and effects; but the vice of an action (its wickedness) is not found in its causes or effects, but is only apparent when we consult the sentiments of the observer. Therefore moral good and evil are not discovered by reason alone.

Hume also attempts in the *Treatise* to establish the other anti-rationalist thesis, that virtue is not the same as reasonableness and vice is not contrary to reason. He gives two arguments for this. The first, very short, argument he claims follows directly from the Representation Argument, whose conclusion was that passions, volitions, and actions can be neither reasonable nor unreasonable. Actions, he observes, *can* be laudable or blamable. Since actions cannot be reasonable or against reason, it follows that “[l]audable and blamable are not the same with reasonable or unreasonable” (T 458). The properties are not identical.

The second and more famous argument makes use of the conclusion defended earlier that reason alone cannot move us to act. As we have seen, reason alone “can never immediately prevent or produce any action by contradicting or approving of it” (T 458). Morality — this argument goes on — influences our passions and actions: we are often impelled to or deterred from action by our opinions of obligation or injustice. Therefore morals cannot be derived from reason alone. This argument is first introduced as showing it impossible “from reason alone... to distinguish betwixt moral good and evil” (T 457) — that is, it is billed as establishing the epistemic thesis. But Hume also says that, like the little direct argument above, it proves that “actions do not derive their merit from a conformity to reason, nor their blame from a contrariety to it” (T458): it is not the reasonableness of an action that makes it good, or its unreasonable that makes it evil.

This argument about motives concludes that moral judgments or evaluations are not the products of reason alone. From this many draw the sweeping conclusion that for Hume moral evaluations are not beliefs or opinions of any kind, but lack all cognitive content. That is, they take the argument to show that Hume holds a non-propositional view of moral evaluations — and indeed, given his sentimentalism, that he is an emotivist: one who holds that moral judgments are meaningless ventings of emotion that can be neither true nor false. Such a reading should be met with caution, however. For Hume, to say that something is not a product of reason alone is not equivalent to saying it is not a truth-evaluable judgment or belief. Hume does not consider all our (propositional) beliefs and opinions to be products of reason; some arise directly from sense perception, for example, and some from sympathy. Also, perhaps there are (propositional) beliefs we acquire via probable reasoning but not by such reasoning *alone*. One possible example is the
On Hume's view, what is a moral evaluation? Four main interpretations have significant textual support. First, as we

6. The Nature of Moral Judgment

inferences frequently in his writings.

he or she may reach some particular moral conclusions by inference from causal, factual premises (stated in terms of

is/ought paragraph that once a person has the moral concepts as the result of prior experience of the moral sentiments,

abstract and causal relations one has discovered that one comes to have the ideas of virtue and vice; one must

premises using reason alone; rather, one requires some input from sentiment. It is not simply by reasoning from the

Other interpreters — the more cognitivist ones — see the paragraph about 'is' and 'ought' as doing none of the above.

Commentators have proposed various interpretations to avoid these difficulties. One approach is to construe 'reason' as

the name of a process or activity, the comparing of ideas (reasoning), and to construe 'morals' as Hume uses it in this

argument to mean the activity of moral discrimination (making a moral distinction). If we understand the terms this way,

the argument can be read not as showing that the faculty of reason (or the beliefs it generates) cannot cause us to

make moral judgments, but rather as showing that the reasoning process (comparing ideas) is distinct from the process

of moral discrimination. This interpretation does not rely on an assumption about the transitivity of causation and is

consistent with Hume's theory of causation.

5. Is and ought

Hume famously closes the section of the Treatise that argues against moral rationalism by observing that other systems

of moral philosophy, proceeding in the ordinary way of reasoning, at some point make an unremarked transition from

premises whose parts are linked only by "is" to conclusions whose parts are linked by "ought" (expressing a new

relation) — a deduction that seems to Hume "altogether inconceivable" (T3.1.1.27). Attention to this transition would

"subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on

the relations of objects, nor is perceiv'd by reason" (ibid.).

Few passages in Hume's work have generated more interpretive controversy.

According to the dominant twentieth-century interpretation, Hume says here that no ought-judgment may be correctly

inferred from a set of premises expressed only in terms of 'is,' and the vulgar systems of morality commit this logical

fallacy. This is usually thought to mean something much more general: that no ethical or indeed evaluative conclusion

whatsoever may be validly inferred from any set of purely factual premises. A number of present-day philosophers,

including R. M. Hare, endorse this putative thesis of logic, calling it "Hume's Law." (As Francis Snare observes, on this

reading Hume must simply assume that no purely factual propositions are themselves evaluative, as he does not argue

for this.) Some interpreters think Hume commits himself here to a non-propositional or noncognitivist view of moral

judgment — the view that moral judgments do not state facts and are not truth-evaluable. (If Hume has already used

the famous argument about the motivational influence of morals to establish noncognitivism, then the is/ought

paragraph may merely draw out a trivial consequence of it. If moral evaluations are merely expressions of feeling

without propositional content, then of course they cannot be inferred from any propositional premises.) Some see the

paragraph as denying ethical realism, excluding values from the domain of facts.

Other interpreters — the more cognitivist ones — see the paragraph about 'is' and 'ought' as doing none of the above.

Some read it as simply providing further support for Hume's extensive argument that moral properties are not

discernible by demonstrative reason, leaving open whether ethical evaluations may be conclusions of cogent probable

arguments. Others interpret it as making a point about the original discovery of virtue and vice, which must involve the

use of sentiment. On this view, one cannot make the initial discovery of moral properties by inference from nonmoral

premises using reason alone; rather, one requires some input from sentiment. It is not simply by reasoning from the

abstract and causal relations one has discovered that one comes to have the ideas of virtue and vice; one must

respond to such information with feelings of approval and disapproval. Note that on this reading it is compatible with the

is/ought paragraph that once a person has the moral concepts as the result of prior experience of the moral sentiments,

he or she may reach some particular moral conclusions by inference from causal, factual premises (stated in terms of

'is') about the effects of character traits on the sentiments of observers. They point out that Hume himself makes such

inferences frequently in his writings.

6. The Nature of Moral Judgment

On Hume's view, what is a moral evaluation? Four main interpretations have significant textual support. First, as we
have seen, the nonpropositional view says that for Hume a moral evaluation does not express any proposition or state any fact; either it gives vent to a feeling, or it is itself a feeling (Flew, Blackburn, Snare, Bricke). (A more refined form of this interpretation allows that moral evaluations have some propositional content, but claims that for Hume their essential feature, as evaluations, is non-propositional.) The subjective description view, by contrast, says that for Hume moral evaluations describe the feelings of the spectator, or the feelings a spectator would have were she to contemplate the trait or action from the common point of view. Often grouped with the latter view is the third, dispositional interpretation, which understands moral evaluations as factual judgments to the effect that the evaluated trait or action is so constituted as to cause feelings of approval or disapproval in a (suitably characterized) spectator (Mackie, in one of his proposals). On the dispositional view, in saying some trait is good we attribute to the trait the dispositional property of being such as to elicit approval. A fourth interpretation distinguishes two psychological states that might be called a moral evaluation: an Occurrent feeling of approval or disapproval (which is not truth-apt), and a moral belief or judgment that is propositional. Versions of this fourth interpretation differ in what they take to be the content of that latter mental state. One version says that the moral judgments, as distinct from the moral feelings, are factual judgments about the moral sentiments (Capaldi). A distinct version, the moral sensing view, treats the moral beliefs as ideas copied from the impressions of approval or disapproval that represent a trait of character or an action as having whatever quality it is that one experiences in feeling the moral sentiment (Cohon). This last view emphasizes Hume’s claim that moral good and evil are like heat, cold, and colors as understood in “modern philosophy,” which are experienced directly by sensation, but about which we form beliefs.

7. Sympathy, and the Nature and Origin of the Moral Sentiments

Our moral evaluations of persons and their character traits, on Hume’s positive view, arise from our sentiments. The virtues and vices are those traits the disinterested contemplation of which produces approval and disapproval, respectively, in whoever contemplates the trait, whether the trait’s possessor or another. These moral sentiments are emotions (in the present-day sense of that term) with a unique phenomenological quality, and also with a special set of causes. They are caused by contemplating the person or action to be evaluated without regard to our self-interest, and from a common or general perspective that compensates for certain likely distortions in the observer’s sympathies, as explained in Section 8. Approval (approbation) is a pleasure, and disapproval (disapprobation) a pain or uneasiness. The moral sentiments are typically calm rather than violent, although they can be intensified by our awareness of the moral responses of others. They are types of pleasure and uneasiness that are associated with the passions of pride and humility, love and hatred: when we feel moral approval of another we tend to love or esteem her, and when we approve a trait of our own we are proud of it. Some interpreters analyze the moral sentiments as themselves forms of these four passions; others argue that Hume’s moral sentiments tend to cause the latter passions. We distinguish which traits are virtuous and which are vicious by means of our feelings of approval and disapproval toward the traits; our approval of actions is derived from approval of the traits we suppose to have given rise to them. We can determine, by observing the various sorts of traits toward which we feel approval, that every such trait — every virtue — has at least one of the following four characteristics: it is either immediately agreeable to the person who has it or to others, or it is useful (advantageous over the longer term) to its possessor or to others. Vices prove to have the parallel features: they are either immediately disagreeable or disadvantageous either to the person who has them or to others. These are not definitions of ‘virtue’ and ‘vice’ but empirical generalizations about the traits as first identified by their effects on the moral sentiments.

In the Treatise Hume details the causes of the moral sentiments, in doing so explaining why agreeable and advantageous traits prove to be the ones that generate approval. He claims that the sentiments of moral approval and disapproval are caused by some of the operations of sympathy, which is not a feeling but rather a psychological mechanism that enables one person to receive by communication the sentiments of another (more or less what we would call empathy today).

Sympathy in general operates as follows. First, observation of the outward expression of another person’s “affection” (feeling or sentiment) in his “countenance and conversation” conveys the idea of his passion into my mind. So does observing the typical cause of a passion: for example, viewing the instruments laid out for another’s surgery will evoke ideas in me of fear and pain. We at all times possess a maximally vivid and forceful impression of ourselves. According to the dispositional interpretation, vivacity of one perception is automatically transferred to those others that are related to it by resemblance, contiguity, and cause and effect. Here resemblance and contiguity are primary. All human beings, regardless of their differences, are similar in bodily structure and in the types and causes of their passions. The person I observe or consider may further resemble me in more specific shared features such as character or nationality. Because of the resemblance and my contiguity to the observed person, the idea of his passion is associated in my mind with my impression of myself, and acquires great vivacity from it. The sole difference between an idea and an impression is the degree of liveliness or vivacity each possesses. So great is this acquired vivacity that the idea of his passion in my mind becomes an impression, and I actually experience the passion. When I come to share in the affections of strangers, and feel pleasure because they are pleased, as I do when I experience an aesthetic enjoyment of a well-designed ship or fertile field that is not my own, my pleasure can only be caused by sympathy (T 2.2.2–8, 3.3.1.7–8). Similarly, Hume observes, when we reflect upon a character or mental quality knowing its tendency either to the benefit or enjoyment of strangers or to their harm or uneasiness, we come to feel enjoyment when the trait is beneficial or agreeable to those strangers, and uneasiness when the trait is harmful or disagreeable to them. This reaction of ours to the tendency of a character trait to affect the sentiments of those with whom we have no special affectionate ties can only be explained by sympathy.
We greatly approve the artificial virtues (justice with respect to property, allegiance to government, and dispositions to obey the laws of nations and the rules of modesty and good manners), which (Hume argues) are inventions contrived solely for the interest of society. We approve them in all times and places, even where our own interest is not at stake, solely for their tendency to benefit the whole society of that time or place. This instance confirms that “the reflecting on the tendency of characters and mental qualities, is sufficient to give us the sentiments of approbation and blame” (T 3.3.1.9). The sympathy-generated pleasure, then, is the moral approbation we feel toward these traits of character. We find the character traits — the causes — agreeable because they are the means to ends we find agreeable as a result of sympathy. Hume extends this analysis to the approval of most of the natural virtues. Those traits of which we approve naturally (without any social contrivance), such as beneficence, clemency, and moderation, also tend to the good of individuals or all of society. So our approval of those can be explained in precisely the same way, via sympathy with the pleasure of those who receive benefit. And since the imagination is more struck by what is particular than by what is general, manifestations of the natural virtues, which directly benefit any individual to whom they are directed, are even more apt to give pleasure via sympathy than are the manifestations of justice, which may harm identifiable individuals in some cases though they contribute to a pattern of action beneficial to society as a whole (T 3.3.1.13).

8. The Common Point of View

As we saw, the moral sentiments are produced by sympathy with those affected by a trait or action. Such sympathetically-acquired feelings are distinct from our self-interested responses, and an individual of discernment learns to distinguish her moral sentiments (which are triggered by contemplating another’s character trait “in general”) from the pleasure or uneasiness she may feel when responding to a trait with reference to her “particular interest,” for example when another’s strength of character makes him a formidable opponent (T 3.1.2.4).

However, the sympathetic transmission of sentiments can vary in effectiveness depending upon the degree of resemblance and contiguity between the observer and the person with whom he sympathizes. I receive the sentiments of someone very much like me or very close to me in time or place far more strongly than I do those of someone unlike me or more remote from me in location or in history. Yet the moral assessments we make do not vary depending upon whether the person we evaluate resembles us in language, sex, or temperament, or is near or far. Indeed, our moral assessments of people remain stable even though our position with respect to them changes over time. Furthermore, sympathy only brings us people’s actual sentiments or what we believe to be their actual sentiments; yet we feel moral approval of character traits that we know produce no real happiness for anyone, because, for example, their possessor is isolated in a prison. To handle these objections to the sympathy theory, and to explain more generally how, on a sentiment-based ethical theory, moral evaluations made by one individual at different times and many individuals in a community tend to be fairly uniform, Hume claims that people do not make their moral judgments from their own individual points of view, but instead select “some common point of view, from which they might survey their object, and which might cause it to appear the same to all of them” (T 3.3.1.30). At least with respect to natural virtues and vices, this common point of view is composed of the intimate perspectives of the various individuals who have direct interactions with the person being evaluated. To make a moral evaluation I must sympathize with each of these persons in their dealings with the subject of my evaluation; the blame or praise I give as a result of this imaginative exercise is my genuine moral assessment of the subject’s character. In that assessment I also overlook the small external accidents of fortune that might render an individual’s trait ineffectual, and respond to traits that render a character typically “fitted to be beneficial to society,” even if circumstances do not permit it to cause that benefit (T 3.3.1.20). Thus I acquire by sympathy the pleasure or uneasiness that I imagine people would feel were the trait able to operate as it ordinarily does. “Experience soon teaches us this method of correcting our sentiments, or at least, of correcting our language, where the sentiments are more stubborn and inalterable” (T 3.3.1.16).

9. Artificial and Natural Virtues

The standard object of moral evaluation is a “quality of mind,” a character trait. (As we have seen, for Hume evaluation of an action is derived from evaluation of the inner quality we suppose to have given rise to it.) The typical moral judgment is that some trait, such as a particular person’s benevolence or laziness, is a virtue or a vice. A character trait, for Hume, is a psychological disposition consisting of a tendency to feel a certain sentiment or combination of sentiments, ones that often move their possessor to action. We reach a moral judgment by feeling approval or disapproval upon contemplating someone’s trait in a disinterested way from the common point of view. So moral approval is a favorable sentiment in the observer elicited by the observed person’s disposition to have certain motivating sentiments. Thus moral approval is a sentiment that is directed toward sentiments, or the dispositions to have them.

In the Treatise Hume emphasizes that “our sense of every kind of virtue is not natural; but ... there are some virtues, that produce pleasure and approbation by means of an artifice or contrivance, which arises from the circumstances and necessities of mankind” (T 3.2.1.1). He divides the virtues into those that are natural — in that our approval of them does not depend upon any cultural inventions or jointly-made social rules — and those that are artificial (dependent both for their existence as character traits and for their ethical merit on the presence of conventional rules for the common good), and he gives separate accounts of the two kinds. The traits he calls natural virtues are more refined and completed forms of those human sentiments we could expect to find even in people who belonged to no society but cooperated only within small familial groups. The traits he calls artificial virtues are the ones we need for successful
impartial cooperation; our natural sentiments are too partial to give rise to these without intervention. In the Treatise Hume includes among the artificial virtues honesty with respect to property (which he often calls equity or “justice,” though it is a strangely narrow use of the term), fidelity to promises (sometimes also listed under “justice”), allegiance to one’s government, conformity to the laws of nations (for princes), chastity (refraining from non-marital sex) and modesty (both primarily for women and girls), and good manners. A great number of individual character traits are listed as natural virtues, but the main types discussed in detail are greatness of mind (“a hearty pride, or self-esteem, if well-concealed and well-founded,” T 3.2.2.11), goodness or benevolence (an umbrella category covering generosity, gratitude, friendship, and more), and such natural abilities as prudence and wit, which, Hume argues, have a reasonably good claim to be included under the title moral virtue, though traditionally they are not. Hume does not explicitly draw a distinction between artificial and natural virtues in the moral Enquiry.

In the Treatise Hume argues in turn that the virtues of material honesty and of faithfulness to promises and contracts are artificial, not natural virtues. Both arguments fall into at least two stages: one to show that if we suppose the given character trait to exist and to win our approval without help from any cooperative social arrangement, paradoxes arise; and another, longer stage to explain how the relevant convention might have come into being and to refute those with a different genetic story. He also explains the social construction of the other artificial virtues and what social good they serve.

10. Honesty with respect to Property

10.1 The Circle

Hume offers a rather cryptic argument to show that our approval of material honesty must be the product of collaborative human effort (convention). When we approve an action, he says, we regard it merely as the sign of the motivating passion in the agent’s “mind and temper” that produced it; our evaluation of the action is derived from our assessment of this inner motive. Therefore all actions deemed virtuous derive their goodness only from virtuous motives — motives we approve. It follows from this that the motive that originally “bestows a merit on any action” can never be moral approval of that action (awareness of its virtue), but must be a non-moral, motivating psychological state — that is, a state distinct from the “regard to the virtue” of an action (moral approval or disapproval) (T 3.2.1.4). For if the virtue-bestowing motive of the action were the agent’s sense that the act would be virtuous to do — if that were why he did it, and why we approved it — then we would be reasoning in a vicious circle: we would approve of the action derivatively, because we approve of the agent’s motive, and this motive would consist of approval of the action, which can only be based on approval of a motive... The basis of our approval could not be specified. For every virtue, therefore, there must be some non-moral motive that characteristically motivates actions expressive of that virtue, which motive, by eliciting our approval, makes the actions so motivated virtuous. The virtue of an action of this species would be established by its being done from this non-moral motive, and only then could an agent also or alternatively be moved so to act by her derivative concern for the virtue of the act. However, Hume observes that there is no morally approved (and so virtue-bestowing), non-moral motive of honest action. The only approved, reliable motive that we can find for acts of “equity” is a moral one, the sense of virtue or “regard to the honesty” of the actions. The honest individual repays a loan not (merely) out of self-interest or concern for the well-being of the lender (who may be a “profligate debaucher” who will reap only harm from his possessions), but from a “regard to justice, and abhorrence of villainy and knavery” (T 3.2.1.9, 13). This, however, is “evident sophistry and reasoning in a circle...” Now nature cannot have “establish’d a sophistry, and render’d it necessary and unavoidable...”; therefore, “the sense of justice and injustice is not deriv’d from nature, but arises artificially... from education, and human conventions” (T 3.2.1.17). Whatever, exactly, the logic of this argument is supposed to be, Hume’s intent is to show that if we imagine equity to be a natural virtue we commit ourselves to a sophistry, and therefore honesty is instead man-made.

Hume offers an account of the genesis of the social convention that creates honesty with respect to property, and this is meant to cope in some way with the circularity he identifies. How it does so is a matter of interpretive controversy, as we will see.

10.2 The Origin of Material Honesty

Hume next poses two questions about the rules of ownership of property and the associated virtue of material honesty: what is the artifice by which human beings create them, and why do we attribute moral goodness and evil to the observance and neglect of these rules?

By nature human beings have many desires but are individually ill-equipped with strength, natural weapons, or natural skills to satisfy them. We can remedy these natural defects by means of social cooperation: shared strength, division of labor, and mutual aid in times of individual weakness. It occurs to people to form a society as a consequence of their experience with the small family groups into which they are born, groups united initially by sexual attraction and familial love, but in time demonstrating the many practical advantages of working together with others. However, in the conditions of moderate scarcity in which we find ourselves, and given the portable nature of the goods we desire, our untrammeled greed and naturally “confined generosity” (generosity to those dear to us in preference to others) tends to create conflict or undermine cooperation, destroying collaborative arrangements among people who are not united by ties of affection, and leaving us all materially poor. No remedy for this natural partiality is to be found in “our natural uncultivated ideas of morality” (T 3.2.2.8); an invention is needed.
Hume argues that we create the rules of ownership of property originally in order to satisfy our avidity for possessions for ourselves and our loved ones, by linking material goods more securely to particular individuals so as to avoid conflict. Within small groups of cooperators, individuals signal to one another a willingness to conform to a simple rule: to refrain from the material goods others come to possess by labor or good fortune, provided those others will observe the same restraint toward them. (This rule will in time require more detail: specific rules determining who may enjoy the creation of ownership does or can depend on any promise or contract, and its concept of convention as an informal practice of mutual compromise for mutual advantage that arises incrementally and entirely informally, without regulating it with rules of justice.

Greed, and more broadly, self-interest, is the motive for inventing property; but we need a further explanation why we think of justice (adherence to the rules of ownership) as virtuous, and injustice (their violation) as vicious. Hume accounts for the moralization of property as follows. As our society grows larger, we may cease to see our own property violations as a threat to the continued existence of a stable economic community, and this reduces our incentive to conform. But when we consider violations by others, we partake by sympathy in the uneasiness these violations cause to their victims and all of society. Such disinterested uneasiness, and the concomitant pleasure we feel on contemplating the public benefits of adherence, are instances of moral disapproval and approval. We extend these feelings to our own behavior as a result of general rules. This process is “forwarded by the artifice of politicians” (T 3.2.2.25), who assist nature by cultivating widespread esteem for justice and abhorrence of injustice in order to govern more easily. Private education assists in this further artifice. Thus material honesty becomes a virtue.

10.3 The Motive of Honest Actions

Does this account resolve the circularity problem? Is there any non-moral motive of honest action? Some interpreters say yes, it is greed redirected, which removes the circle. But this presents two difficulties: first, our greed is not in fact best satisfied by just action in every case, and second, Hume denies that this motive is approved. Some interpret Hume as coping with the first difficulty by supposing that politicians and parents deceive us into thinking, falsely, that every individual just act advances the interests of the agent; or they claim that Hume himself mistakenly thought so, at least in the *Treatise* (see Baron, Haakonssen, and Gauthier). Others claim that Hume identifies a non-moral motive of honest action (albeit an artificial one) other than redirected greed, such as a disposition to treat the rules of justice as themselves reason-giving (Darwall) or having a policy of conforming to the rules of justice as a system (Garrett). Still others say there is no non-moral motive of honest action, and Hume escapes from the circle by relaxing this ostensibly universal requirement on virtuous types of behavior, limiting it to the naturally virtuous kinds. These interpreters either claim that there is no particular motive needed to evoke approval for conformity to the rules of property — mere behavior is enough (Mackie) — or that we approve of a motivating form of the moral sentiment itself, the sense of duty (Cohon).

Hume’s genetic account of property is striking for its lack of patriarchal assumptions about the family, its explicit denial that the creation of ownership does or can depend on any promise or contract, and its concept of convention as an informal practice of mutual compromise for mutual advantage that arises incrementally and entirely informally, without the use of central authority or force.

11. Fidelity to Promises

Fidelity is the virtue of being disposed to fulfill promises and contracts. Hume has in mind promises made “at arm’s length” that parties undertake to promote their own interest, not affectionate exchanges of favors between friends. While he identifies the same circularity puzzle about the approved motive of fidelity that he tackles at length in connection with honesty, in the case of fidelity he concentrates on a different conundrum that arises with the misguided attempt to analyze fidelity as a non-conventional (natural) virtue. Unlike Hobbes and Locke, who help themselves to the concept of a promise or contract in their imagined state of nature, Hume argues that the performative utterance “I promise” would be unintelligible in the absence of background social conventions, and that the moral obligation of a promise is dependent upon such conventions as well.

Suppose the practice of giving and receiving promises did not depend on a socially-defined convention. In that case, what could we mean by the utterances we use to make them, and what would be the origin of our obligation to fulfill them? Where the words are used (uncharacteristically) in a way that does not purport to reveal the agent’s will (as when the person is joking or play-acting), we do not think a promise is really being made; we only take a speaker to have promised, and so to be bound to perform, if he understands the words he uses, in particular as purporting to oblige him. Thus for effective use there must be some act of the speaker’s mind expressed by the special phrase “I promise” and its synonyms, and our moral obligation results from this act of the mind. (This seems to be Hobbes’s assumption in
Leviathan, where the implicit signs of covenant — as distinct from the explicit ones — are clear signs of the person’s will.) The requisite mental act or mental state, though, could not be one of mere desire or resolution to act, since it does not follow from our desiring or resolving to act that we are morally obligated to do so; nor could it be the volition to act, since that does not come into being ahead of time when we promise, but only when the time comes to act. And of course, one can promise successfully (incur obligation by promising) even though one has no intention to perform; so the mental act requisite to obligation is not the intention to perform. The only likely act of mind that might be expressed in a promise is a mental act of willing to be obligated to perform the promised action, as this conforms to our common view that we bind ourselves by choosing to be bound.

But, Hume argues, it is absurd to think that one can actually bring an obligation into existence by willing to be obligated. What makes an action obligatory is that its omission is disapproved by unbiased observers. But no act of will within an agent can directly change a previously neutral act into one that provokes moral disapproval in observers (even in the agent herself). Sentiments are not subject to such voluntary control. Even on a moral rationalist view the thesis would be absurd: to create a new obligation would be to change the abstract relations in which actions and persons stand to one another, and one cannot do this by performing in one’s own mind an act of willing such a relation to exist. Thus, there is no such act of the mind. Even if people in their natural (pre-conventional) condition “cou’d perceive each other’s thoughts by intuition,” they could not understand one another to bind themselves by any act of promising, and could not be obligated thereby. Since the necessary condition for a natural obligation of promises cannot be fulfilled, we may conclude that this obligation is instead the product of group invention to serve the interests of society.

Promises are invented in order to build upon the advantages afforded by property. The invention of mere ownership suffices to make possession stable. The introduction of transfer by consent permits some trade, but so far only simultaneous swapping of visible commodities. Great advantages could be gained by all if people could be counted on to provide goods or services later for benefits given now, or exchange goods that are distant or described generically. But for people without the capacity to obligate themselves to future action, such exchanges would depend upon the party who performs second doing so out of gratitude alone; and that motive cannot generally be relied on in self-interested transactions. However, we can devise better ways to satisfy our appetites “in an oblique and artificial manner...” (T 3.2.5.9). First, people can easily recognize that additional kinds of mutual exchanges would serve their interests. They need only express this interest to one another in order to encourage everyone to invent and to keep such agreements. They devise a form of words to mark these new sorts of exchanges (and distinguish them from the generous reciprocal acts of friendship and gratitude). When someone utters this form of words, he is understood to express a resolution to do the action in question, and he “subjects himself to the penalty of never being trusted again in case of failure” (T 3.2.5.10), a penalty made possible by the practice of the group, who enforce the requirement to keep promises by the simple expedient of refusing to contract with those whose word cannot be trusted. This “concert or convention” (ibid.) alters human motives to act. One is moved by self-interest to give the promising sign (in order to obtain the other party’s cooperation), and once one has given it, self-interest demands that one do what one promised to do so as to insure that people will exchange promises with one in the future. Some interpreters say that this enlightened self-interest remains the only motive for keeping one’s promise, once the practice of promising has been created. But Hume says the sentiment of morals comes to play the same role in promise-keeping that it does in the development of honesty with respect to property (T 3.2.5.12); so there is evidence he thinks the moral sentiment not only becomes “annex’d” to promise-keeping but further motivates it. In larger, more anonymous communities, a further incentive is needed besides the fear of exclusion; and a sentiment of moral approval of promise-keeping arises as the result of sympathy with all who benefit from the practice, aided by a “second artifice,” the well-meaning psychological manipulation of the people by parents and politicians, which yields a near-universal admiration of fidelity and shame at breaking one’s word (T 3.2.5.12). This may provide a moral motive for promise-keeping even in anonymous transactions.

12. Allegiance to Government

A small society can maintain a subsistence-level economy without any dominion of some people over others, relying entirely on voluntary compliance with conventions of ownership, transfer of goods, and keeping of agreements, and relying on exclusion as the sole means of enforcement. But an increase in population and/or material productivity, Hume thinks, tends to stimulate a destabilizing rate of defection from the rules: more luxury goods greatly increase the temptation to act unjustly, and more anonymous transactions make it seem likely that one will get away with it. Though people are aware that injustice is destructive of social cooperation and ultimately detrimental to their own interests, this knowledge will not enable them to resist such strong temptation, because of an inherent human weakness: we are more powerfully drawn to a near-term good even when we know we will pay for it with the loss of a greater long-term good. This creates the need for government to enforce the rules of property and promise (the “laws of nature,” as Hume sometimes rather ironically calls them, since on his view they are not natural). This is the reason for the invention of government. Once in power, rulers can also make legitimate use of their authority to resolve disputes over just what the rules of justice require in particular cases, and to carry out projects for the common good such as building roads and dredging harbors.

Hume thinks it unnecessary to prove that allegiance to government is the product of convention and not mere nature, since governments are obviously social creations. But he does need to explain the creation of governments and how they solve the problem he describes. He speculates that people who are unaccustomed to subordination in daily life might draw the idea for government from their experience of wars with other societies, when they must appoint a
especially clear with such self-regarding virtues as prudence and industry, which we approve even when they occur in
into our minds the pleasure that the trait produces for its possessor or for others (with one minor exception). This is
possessor or to others. In these four groups of approved traits, our approval arises as the result of sympathy bringing
fall into four groups: traits immediately agreeable to their possessor or to others, and traits advantageous to their
action when we consider it from the common point of view. As we saw, he argues that the traits of which we approve
such natural virtues as gratitude and friendship because we sympathize with those who are affected by each such
vices. Sympathy also explains our approval of the artificial virtues; the difference is that we approve of those as a result
The mechanism of sympathy ultimately accounts for this approval and the corresponding disapproval of the natural
of government; so our duty of allegiance forbids this. A government that maintains conditions preferable to what they
would be without it retains its legitimacy and may not rightly be overthrown. But rebellion against a cruel tyranny is no
violation of our duty of allegiance, and may rightly be undertaken.

Hume does advocate some forms of government as being preferable to others, particularly in his Essays. Governments
structured by laws are superior to those controlled by the edicts of rulers or ruling bodies (“That Politics May Be
Reduced to a Science”). Representative democracy is superior to direct democracy, and “free” (popular) governments
are more hospitable to trade than “absolute” governments (ibid.). Hume speculates that a perfect government would be
a representative democracy of property-holders with division of powers and some features of federalism (“Idea of a
Perfect Commonwealth”). He defends his preferences by arguing that certain forms of government are less prone to
corruption, faction (with the concomitant threat of civil war), and oppressive treatment of the people than others; that is,
they are more likely to enforce the rules of justice, adjudicate fairly, and encourage peace and prosperity.

Hume famously criticizes the social contract theory of political obligation. According to his own theory, our duty to obey
our governors is not reducible to an instance of our duty to fulfill promises, but arises separately though in a way
parallel to the genesis of that duty. Hume denies that any native citizen or subject in his own day has made even a tacit
promise to obey the government, given that citizens do not think they did any such thing, but rather think they are born
to obey it. Even a tacit contract requires that the will be engaged, and we have no memory of this; nor do governments
refrain from punishing disloyalty in citizens who have given no tacit promise.

13. The Natural Virtues

In the Treatise Hume’s principle interest in the natural virtues lies in explaining the causes that make us approve them.
The mechanism of sympathy ultimately accounts for this approval and the corresponding disapproval of the natural
vices. Sympathy also explains our approval of the artificial virtues; the difference is that we approve of those as a result
of sympathy with the cumulative effects produced by the general practice of the artificial virtues on the whole of society
(individual acts of justice not always producing pleasure for anyone); whereas we approve each individual exercise of
such natural virtues as gratitude and friendship because we sympathize with those who are affected by each such
action when we consider it from the common point of view. As we saw, he argues that the traits of which we approve
fall into four groups: traits immediately agreeable to their possessor or to others, and traits advantageous to their
possessor or to others. In these four groups of approved traits, our approval arises as the result of sympathy bringing
into our minds the pleasure that the trait produces for its possessor or for others (with one minor exception). This is
especially clear with such self-regarding virtues as prudence and industry, which we approve even when they occur in
According to Hume, different levels and manifestations of the passions of pride and humility make for virtue or for vice. An obvious and "over-weaning conceit" is disapproved by any observer (is a vice) (T 3.3.2.1); while a well-founded but concealed self-esteem is approved (is a virtue). Hume explains these opposite reactions to such closely related character traits by means of the interplay of the observer's sympathy with a distinct psychological mechanism he calls comparison. The mechanism of comparison juxtaposes a sympathetically-communicated sentiment with the observer's own inherent feeling, causing the observer to feel a sentiment opposite to the one she observes in another (pleasure if the other is suffering, pain if the other is pleased) when the sympathetically-communicated sentiment is not too strong. A person who displays excessive pride irritates others because, while others come to feel this person's pleasant sentiment of pride (to some degree) via sympathy, they also feel a greater uneasiness as a result of comparing that great pride (in whose objects they do not believe) with their own lesser pride in themselves; this is why conceit is a vice. Self-esteem founded on an accurate assessment of one's strengths and politely concealed from others, though, is both agreeable and advantageous to its possessor without being distressing to others, and so is generally approved. (Thus the professed preference of Christians for humility over self-esteem does not accord with the judgments of most observers.) Although excessive pride is a natural vice and self-esteem a natural virtue, human beings in society create the artificial virtue of good breeding (adherence to customs of slightly exaggerated mutual deference in accordance with social rank) to enable us each to conceal our own pride easily so that it does not shock the pride of others.

Courage and military heroism are also forms of pride. Though the student of history can see that military ambition has mostly been disadvantageous to human society, when we contemplate the "dazzling" character of the hero, immediate sympathy irresistibly leads us to approve it (T 3.3.2.15).

Our approval of those traits may be grouped together under the heading of goodness and benevolence, such as generosity, humanity, compassion, and gratitude, arises from sympathy with people in the individual's "narrow circle" of friends and associates, since, given natural human selfishness, we cannot expect people's concerns to extend farther (T 3.3.3.2). By adopting the common point of view we correct for the distortions of sympathy by entering into the feelings of those close to the person being evaluated even if they are remote from us. The vice of cruelty is mostly loathed because the suffering of the person's victims that reaches us via sympathy readily becomes hatred of the perpetrator.

Although natural abilities of the mind are not traditionally classified as moral virtues and vices, the difference between these types of traits is unimportant, Hume argues. Intelligence, good judgment, application, eloquence, and wit are also mental qualities that bring individuals the approbation of others, and their absence is disapproved. As is the case with many of the traditionally-recognized virtues, the various natural abilities are approved either because they are useful to their possessor or because they are immediately agreeable to others. It is sometimes argued that moral virtues are unlike natural abilities in that the latter are involuntary, but Hume argues that many traditional moral virtues are involuntary as well. The sole difference is that the prospect of reward or punishment can induce people to act as the morally virtuous would (as justice requires, for example), but cannot induce them to act as if they had the natural abilities.

14. Differences between the Treatise and the Moral Enquiry

Late in his life Hume deemed the Enquiry concerning the Principles of Morals his best work, and in style it is a model of elegance and subtlety. His method in that work differs from that of the Treatise: instead of explicating the nature of virtue and vice and our knowledge of them in terms of underlying features of the human mind, he proposes to collect all the traits we know from common sense to be virtues and vices, observe what those in each group have in common, and from that observation discover the "foundation of ethics" (EPM 1.10). The conclusions largely coincide with those of the Treatise. Some topics in the Treatise are handled more fully in the moral Enquiry; for example Hume's account of the motive to just action is enriched by his discussion of a challenge from a "sensible knave." However, without the detailed background theories of the mind, the passions, motivation to action, and social convention presented in the Treatise, and without any substitute for them, some of the conclusions of the moral Enquiry stand unsupported.

In the latter work, Hume's main argument that reason alone is not adequate to yield moral evaluations (in Appendix 1) depends on his having demonstrated throughout the book that at least one foundation of moral praise lies in the usefulness to society of the praised character trait. We use reason extensively to learn the effects of various traits and to identify the useful and pernicious ones. But utility and disutility are merely means; were we indifferent to the weal and woe of mankind, we would feel equally indifferent to the traits that promote those ends. Therefore there must be some sentiment that makes us favor the one over the other. This could only be humanity, "a feeling for the happiness of mankind, and resentment of their misery" (EPM App. 1.3). This argument presupposes that the moral evaluations we make are themselves the expression of sentiment rather than reason alone. (The alternative position would be that while of course we do feel approval and disapproval for vice and virtue, the judgment as to which is which is itself the deliverance of reason.) So Hume appends some arguments directed against the hypothesis of moral rationalism. One of these is an enriched version of the argument of Treatise 3.1.1 that neither demonstrative nor causal reasoning has moral distinctions as its proper object, since moral vice and virtue cannot plausibly be analyzed as either facts or relations. He adds that while in our reasonings we start from the knowledge of relations or facts and infer some previously-unknown relation or fact, moral evaluation cannot proceed until all the relevant facts and relations are
already known. At that point, there is nothing further for reason to do; therefore moral evaluation is not the work of reason alone but of another faculty. He bolsters this line of argument by expanding his Treatise analogy between moral and aesthetic judgment, arguing that just as our appreciation of beauty awaits full information about the object but requires the further contribution of taste, so in moral evaluation our assessment of merit or villainy awaits full knowledge of the person and situation but requires the further contribution of approbation or disapprobation. He also offers the argument that since the chain of reasons why one acts must finally stop at something that is "desirable on its own account... because of its immediate accord or agreement with sentiment..." (EPM App.1.19), sentiment is needed to account for ultimate human ends; and since virtue is an end, sentiment and not reason alone must distinguish moral good and evil.

In the moral Enquiry Hume omits all arguments to show that reason alone does not move us to act; so the Representation Argument about the irrelevance of reason to passions and actions is absent. Without it he has no support for his direct argument that moral goodness and evil are not identical with reasonableness and unreasonableness, which relies on it for its key premise; and that too is absent from EPM. On the whole in EPM Hume does not appeal to the thesis that reason cannot produce motives in order to show that morals are not derived from reason alone, but limits himself to the epistemic and descriptive arguments showing that reason alone cannot discern virtue and vice in order to reject ethical rationalism in favor of sentimentalism. However, at Appendix I.21 he does assert (without support) that "Reason, being cool and disengaged, is no motive to action," and perhaps this is intended to be a premise in a revised version of the famous argument that reason cannot produce motives but morals can, though what he writes here is tantalizingly different from that argument as it appears repeatedly in the Treatise.

Why did Hume omit the more fundamental arguments for the motivational inertia of reason? He may have reconsidered and rejected them. For example, he may have given up his undefended claim that passions have no representative character, a premise of the Representation Argument on which, as we saw, some of his fundamental anti-rationalist arguments depend. Or he may have retained these views but opted not to appeal to anything so arcane in a work aimed at a broader audience and intended to be as accessible as possible. The moral Enquiry makes no use of ideas and impressions, and so no arguments that depend on that distinction can be offered there, including the Representation Argument. Apparently Hume thought he could show that reason and sentiment rule different domains without using those arguments.

Thus, not surprisingly, the causal analysis of sympathy as a mechanism of vivacity-transferal from the impression of the self to the ideas of the sentiments of others is entirely omitted from the moral Enquiry. Hume still appeals to sympathy there to explain the origin of all moral approval and disapproval, but he explains our sympathy with others simply as a manifestation of the sentiment of humanity, which is given more prominence. He is still concerned about the objection that sympathetically-acquired sentiments vary with spatial and temporal distance from the object of evaluation while moral assessments do not; so he addresses it in the moral Enquiry as well, and resolves it by appealing once again to the common point of view. In the Enquiry he places more emphasis on sympathy with the interests of the whole of society, in part achieved by conversation using shared moral vocabulary, as a way to correct our initial sentiments to make them genuinely moral (Taylor 2002). He also attends more explicitly to the role of reason and reflection in moral evaluation. Some interpreters see him as offering an account of how to arrive at reliable moral judgment superior to that in the Treatise (Taylor 2015).

The distinction between artificial and natural virtues that dominates the virtue ethics of the Treatise is almost entirely absent from the moral Enquiry; the term ‘artificial’ occurs in the latter only once in a footnote. Gone are the paradoxes of property and promises intended to prove that particular virtues are devised on purpose; also missing is what some commentators think Hume’s most original contribution to the theory of justice, his account of convention. Yet Hume briefly sketches part of the same quasi-historical account of the origin of justice that he gives in the Treatise; and while the emphasis has shifted, Hume not only tries to show that justice has merit only because of its beneficial consequences, but that “public utility is the sole origin of justice” — were we not to find it useful (and in some conditions we might not) we would not even have such a thing (EPH 3.1.1). While any explanation of this shift and these omissions is merely speculative, here it seems that Hume does not change his mind about the arguments of the Treatise but chooses to lead the reader to the same conclusions by more subtle and indirect means while avoiding provocative claims.

In the moral Enquiry Hume is more explicit about what he takes to be the errors of Christian (or, more cautiously, Roman Catholic) moralists. Not only have they elevated craven humility to the status of a virtue, which he hints in the Treatise is a mistake, but they also favor penance, fasting, and other “monkish virtues” that are in fact disapproved by all reasonable folk for their uselessness and disagreeableness, and so are in fact vices.

Bibliography

Primary Sources: Works by Hume


**Academic Tools**

- How to cite this entry.
- Preview the PDF version of this entry at the Friends of the SEP Society.
- Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
- Enhanced bibliography for this entry at PhilPapers, with links to its database.

**Other Internet Resources**

- The Hume Society
- “Moral Theory”, section of the entry on Hume, by James Fieser (U. Tennessee/Martin), in the *Internet Encyclopedia of Philosophy*.

**Related Entries**

Hume, David | property and ownership

---

*Permission to reproduce material from this entry is granted to academic institutions for non-commercial, educational purposes provided the source is acknowledged. For other uses, contact the SEP.*

*The Encyclopedia Now Needs Your Support*

Please Read How You Can Help Keep the Encyclopedia Free
This is Hume’s principle philosophical work, the central notions of which were rewritten more popularly in Philosophical Essays Concerning Human Understanding (1748) and An Enquiry Concerning the Principles of Morals (1751). This thirty-four page pamphlet was published anonymously and was prompted by Hume’s candidacy for the Chair of Moral Philosophy at the University of Edinburgh. The pamphlet responds to criticisms regarding the Treatise. Philosophical Essays Concerning Human Understanding.