

HUME ON REPRESENTATION, REASON AND MOTIVATION

RACHEL COHON

*Philosophy Department,
Stanford University,
STANFORD CA 94305-2155,
U.S.A.*

cohon@leland.stanford.edu

DAVID OWEN

*Philosophy Department,
213 Social Science Building, # 27,
University of Arizona,
TUCSON AZ 85721,
U.S.A.*

dwo@u.arizona.edu

In this paper, we re-evaluate some of Hume's most famous arguments in Book II and III in the light of some of the positions he holds in Book I of the "Treatise". Hume's Representation Argument in Book II holds that a passion is "an original existence... and contains not any representative quality" (T 415). We evaluate this claim against the backdrop of Hume's general account of impressions and ideas as found in Book I, and argue that Hume holds that no impressions represent: representation is a function limited to ideas. We then examine the role the Representation Argument plays in Hume's arguments, in Books II and III, about the inertia of reason with respect to motivation. Our interpretation of these arguments is constrained by Book I's "Of the influence of belief", where Hume argues that it is a distinguishing characteristic of at least some beliefs that they do indeed motivate. The resulting interpretations hold no comfort for noncognitivist readings of Hume.

PART ONE: INTRODUCTION¹

In a well known passage, Hume says:

A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high. 'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason; since this contradiction consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent. (T 415)

The passage occurs in Book 2, Part 3, Section 3, "Of the influencing motives of the will." Let us call it "The Representation Argument." Very roughly, the argument maintains that since passions have no representative function, they cannot be opposed to or by reason. The same argument, slightly enlarged to include actions and volitions as well as passions, occurs 43 pages later, in Book 3, Part 1, section 1, "Moral distinctions not deriv'd from reason."² Hume there says it serves two purposes.

¹ For comments and advice on an earlier version of this paper, we wish to thank Elijah Millgram, David Fate Norton, Annette Baier and Don Garrett. An earlier version was read at the Hume Society conference in Nottingham, England, July 1996. Geoffrey Sayre-McCord's reply was most helpful. Although we have made some changes in light of his reply, we have tried to keep the main arguments intact, so as not to move the target after the shot had been fired.

² "Reason is the discovery of truth or falsehood. Truth or falsehood consists in an agreement or disagreement either to the real relations of ideas, or to real existence and matter of fact. Whatever, therefore, is not susceptible of this agreement or disagreement, is incapable of being true or false, and can never be an object of our rea-

It proves directly "that actions do not derive their merit from a conformity to reason ...; and it proves the same truth more *indirectly*, by shewing us, that as reason can never immediately prevent or produce any action by contradicting or approving of it, it cannot be the source of the distinction betwixt moral good and evil, which are found to have that influence." (T 458)

In spite of the obvious centrality of The Representation Argument, and its apparently clear and unequivocal expression in two places, there are at least two *prima facie* problems in taking it at face value. (1) The first premise seems to be in conflict with the fact that the passions apparently do represent things to us.³ Anger, according to the account developed by Hume in Book 2, is typically directed; the blind, undirected anger mentioned in the above formulation of The Representation Argument seems atypical, degenerate and in need of explanation with reference to the more typical, directed, central case.⁴ Given this difficulty, why does Hume claim that passions

son. Now 'tis evident our passions, volitions, and actions are not susceptible of any such agreement or disagreement; being original facts and realities, compleat in themselves, and implying no reference to other passions, volitions, and actions. 'Tis impossible, therefore, they can be pronounced either true or false, and be either contrary or conformable to reason." (T 458)

³ As Annette Baier has reminded us. See Baier (1991), pp. 157 ff. For criticism, see Rachel Cohon (1994), and Elijah Millgram (1995).

⁴ Hence Baier's description of the first formulation of the Representation Argument as "one very silly paragraph" (p. 160). As David Norton reminded us, this problem concerns the intentionality of anger, and it is not at all obvious that this is the very same issue as "being a copy of any other existence." However the passage from T 415 seems to show that Hume took intentionality or "having a reference to any other object" as at least a necessary condition of a per-

do not represent? (2) The Representation Argument's indirect role, in helping to show that reason, or any conclusion of reason, doesn't motivate (that being the unique role of the passions), seems to be in tension with the important Book 1, Part 3, Section 10, "Of the influence of belief." There, an important feature of beliefs, as opposed to the merely conceived ideas of the imagination, is that beliefs have an influence on passions and actions.⁵ Why, then, does Hume say that reason doesn't motivate? In this paper, we hope to solve these problems. First, we will establish that in Book 1 Hume shows that impressions do not represent; that is the role of ideas. The main premise of the Representation Argument will then be seen as, at least in part, just a consequence of this general truth. We will then offer an explanation of how Hume can account for the directedness of the passions, in spite of the fact that they are "original existences." In light of this account, we will clarify the relation between reason and motivation, and make some general claims about the scope of the claim that moral distinctions are not based on reason.

PART TWO: IMPRESSIONS, IDEAS AND REPRESENTATION

Locke characterized the term 'idea' as "whatsoever is the Object of the Understanding when a Man thinks"⁶. This is general enough, but he also thinks of many ideas, including ideas of sense, as being essentially representational, serving as signs for something beyond themselves. For instance he says

ception containing a "representative quality, which renders it a copy of any other existence." See also Cohon (1994).

⁵ See also Book 1, Part 3, section 16, "Of the reason of animals," where Hume is happy to talk, not just of beliefs affecting actions, but even of actions which "proceed from a reasoning" (T 177).

⁶ John Locke (1975), p. 47 (1.1.8).

"'Tis therefore the actual receiving of *Ideas* from without, that gives us notice of the Existence of other Things, and makes us know, that something doth exist at that time without us, which causes that *Idea* in us" ((1975), p. 630 (4.11.2)). Ideas, as received in sensation, are for Locke essentially representational: "since the Things, the Mind contemplates, are none of them, besides it self, present to the Understanding, 'tis necessary that something else, as a Sign or Representation of the thing it considers, should be present to it: and these are Ideas." ((1975), pp. 720-721 (4.21.4))

By contrast, Hume, from the very first sentence of the *Treatise*, distinguishes all "the perceptions of the human mind ... into two distinct kinds": impressions and ideas (T1). Implicitly in the *Treatise*, and explicitly in the first *Enquiry*, Hume argues that by missing this distinction, Locke missed the point about the innateness controversy. (E 22, n. 1) Once it is established that all ideas are derived from impressions, it follows, if by 'innate' we mean "original or copied from no precedent perception," that "all our impressions are innate, and our ideas not innate." (E 22) That is to say, impressions are "original or copied from no precedent perception", while "ideas are preceded by other more lively perceptions, from which they are derived, and which they represent." (T 7)⁷ To say that impres-

⁷ Hume makes the same point in the *Abstract*: "This proposition seems to be equivalent to that which Mr. Locke has taken such pains to establish, viz. That no ideas are innate. Only it may be observed, as an inaccuracy of that famous philosopher, that he comprehends all our perceptions under the term idea, in which sense it is false, that we have no innate ideas. For it is evident our stronger perceptions or impressions are innate, and that natural affection, love of virtue, resentment, and all other passions, arise immediately from nature". (T 648)

sions are not copies of other, precedent perceptions is not to deny that they have causes. Nor is it yet to deny that they might resemble, copy or represent their causes.⁸ But it certainly sets the stage for such a denial, and at the very least should prepare us for the claim that a "passion is an original existence" (*T* 415).

Let us turn to Hume's discussion of how it is that one class of perceptions, ideas, can be derived from, resemble, and represent another class, impressions. One of the more important points established by Hume in the very first section of the *Treatise* is the Priority Principle: "That all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent." (*T* 4) This important claim, which Hume describes as "the first principle ... in the science of human nature" (*T* 7), is explicitly, if not solely, about representation. The representation of impressions by ideas is mentioned frequently throughout this section, and it is not there suggested that impressions themselves might represent something else.

The Priority Principle is in part made up of the Correspondence Rule: there is a one to one correspondence between simple ideas and impressions.⁹ Hume establishes this

⁸ At one point Hume explicitly says that whether perceptions are caused by and resemble external objects is a question to be decided by experience: "It is a question of fact, whether the perceptions of the senses be produced by external objects, resembling them: how shall this question be determined? By experience surely; as all other questions of a like nature. But here experience is, and must be entirely silent." (*E* 153)

⁹ "I venture to affirm, that the rule here holds without any exception, and that every simple idea has a simple impression, which resembles it, and every simple impression a correspondent idea." (*T* 3) Note that some delicate type/token distinctions must be made for

rule of correspondence between simple ideas and impressions by simple observation, and realizes "'tis impossible to prove by a particular enumeration." Instead, he issues a challenge to anyone denying the rule to come up with a counterexample.¹⁰ He then turns to his main task: the tracing of the connections, especially the causal connections, between impressions and ideas. Hume says that "The full examination of this question is the subject of the present treatise" (T 4), an important claim. His method in the application of the science of human nature to the subjects treated in this work will be the tracing of connections between impressions and ideas. But at this early stage, Hume is concerned only to establish the Priority Principle. Hume's attempt comes in two stages, and the argument is summed up as follows:

The constant conjunction of our resembling perceptions, is a convincing proof, that the one are the causes of the other; and this priority of the impressions is an equal proof, that our impressions are the causes of our ideas, not our ideas of our impressions. (T 5)

Simple impressions and ideas come in resembling pairs, and as the former cause the latter, the latter represent the former. Ideas represent, and what they represent is impressions. Impressions don't, it appears, represent at all. At least, they don't represent other perceptions of the mind. If they are to represent something else (external objects, for example),

this rule to be even remotely plausible.

¹⁰ It is sometimes held that by providing the example of the missing shade of blue (T 5-6), Hume meets his own challenge. However note that that example is explicitly held by Hume to be a "contradictory phaenomenon," not to the correspondence rule, but to the principle of priority of impressions over ideas.

and if they are to represent these things in the way ideas represent impressions, then at least two things must be true: (1) they must be caused by these other things, and (2) they must resemble, perhaps even be copies of, these causes. But Hume shows little interest in the causes of impressions, at least of the impressions of sensations. At 1.3.5 he says:

As to those *impressions*, which arise from the senses, their ultimate cause is, in my opinion, perfectly inexplicable by human reason, and 'twill always be impossible to decide with certainty, whether they arise immediately from the objects, or are produc'd by the creative power of the mind, or are deriv'd from the author of our being.

(T84)¹¹

And at 2.1.1, he says:

Original impressions or impressions of sensation are such as without any antecedent perception arise in the soul, from the constitution of the body, from the animal spirits, or from the application of objects to the external organs. (T275)

There is a host of possibilities for the causes of sensation: Hume has ruled out only that they are other perceptions of the mind (that is why he here calls them "original impressions"). But we will never be in a position to know what in fact their causes are.

¹¹ Hume goes to say: "Nor is such a question any way material to our present purpose. We may draw inferences from the coherence of our perceptions, whether they be true or false; whether they represent nature justly or be mere illusions of the senses." This may be thought to provide evidence for the view that Hume still regards the representational nature of impressions as an open question. But it is just as easily interpreted as claiming only that nothing important hangs on the question at this stage.

One of the reasons Hume is reluctant to pursue the question of the causes of impressions of sensation is that he knows, and perhaps knows in advance, that we are not going to get what we want, if what we want as causes are external objects that resemble the impressions they occasion. This is clear at least as early as 1.2.6:

Now since nothing is ever present to the mind but perceptions, and since all ideas are deriv'd from something antecedently present to the mind; it follows, that 'tis impossible for us so much as to conceive or form an idea of any thing specifically different from ideas and impressions. (T 67)

We *think* we can conceive of external objects different from but resembling the impressions of sensation they cause, but this is a mistake. The content of all our ideas comes from impressions. We cannot even conceive what it would be like for a perception of the mind to resemble something that is not a perception of the mind. To paraphrase Berkeley, nothing can be like a perception of the mind but another perception of the mind. Impressions and ideas resemble each other. But ideas are derived from impressions, not the other way around. So ideas can represent, but impressions cannot.

It remains true of course that we (or at least philosophers) believe that impressions of sensation are caused by and resemble, and hence represent, external objects. And Hume himself sometimes speaks that way, e.g. at T 67 and 84.¹² Ordi-

¹² And, at T 38, Hume speaks of that "compound impression, which represents extension". Hume is here saying that our idea of extension is derived from an impression, not that that impression is derived from and resembles some external phenomenon which we call "extension". We thank Don Garrett for pressing us on this and many other examples concerning Hume, impressions, and representations.

nary people have the simpler belief that impressions simply are external objects. Both beliefs, especially the philosopher's, must be explained. And this Hume attempts to do in 1.4.2. The difficult details of Hume's explanation are beyond the scope of this paper; but it is worth quickly summarizing his reasons for thinking that, as is predictable from 1.2.6, the philosopher's belief in distinct existence comes not from sense, nor is it founded on reason. For sense to "produce the opinion of a distinct existence" which causes the impressions which resemble it, or to "offer it to the mind as represented," it must "present both an object and an image." (T 191-93). But this is an impossibility. Nor can we reason from the existence of an impression to a belief in a distinct object which causes it:

But as no beings are ever present to the mind but perceptions; it follows that we may observe a conjunction or a relation of cause and effect between different perceptions, but can never observe it between perceptions and objects. 'Tis impossible, therefore, that from the existence or any of the qualities of the former, we can ever form any conclusion concerning the existence of the latter, or ever satisfy our reason in this particular. (T212)

Indeed, Hume describes the philosophical hypothesis "of the double existence of perceptions and objects" as a "monstrous offspring" (T215).

Some perceptions of the mind, ideas, represent other perceptions, impressions. Contrary to our initial inclination to believe otherwise, it turns out that impressions, at least impressions of sensation, don't represent. So when Hume says, "A passion is an original existence ... and contains not any representative quality" (T 415), we should hardly be surprised. Let us look briefly at the differences between impressions of sensation and impressions of reflection. At 1.1.2, Hume says that an

impression of sensation "arises in the soul originally, from unknown causes." (T 7) But an impression of reflection "is derived in a great measure from ideas" (T 7). We receive an impression of pleasure or pain, from which an idea is taken. And when this idea "returns upon the soul, [it] produces the new impression of desire and aversion, hope and fear, which may properly be called impressions of reflexion, because derived from it." (T 8) Much the same division is made at 2.1.1, where the main division is between original and secondary impressions. The former include "the impressions of the senses, and all bodily pains and pleasures", while the latter include "the passions, and other emotions resembling them." (T 275) Hume's use of the terms "original" and "secondary" are instructive here. By "original", Hume means something like: part of the bedrock of human nature, which can't be explained by appeal to prior causes.¹³ Hume doesn't deny that they have "natural and physical causes" (T 275), but only that such causes, presumably having to do with animal spirits etc., are irrelevant to their nature as original impressions. Secondary impressions, by contrast, have as their cause an idea or another impression. And this causal story is part of what it is for these impressions to be impressions of reflection.

So, when Hume says that a "passion is an original existence" (T 415), he is not using "original" in the same sense as "original (vs. secondary) impression". In the latter case, "original" means something like "from unknown causes". Passions are secondary, not original impressions. Rather, when Hume says that passions are original existences he means that, even though we know the

¹³ See for instance T 280: "Now these qualities, which we must consider as original, are such as are most inseparable from the soul, and can be resolv'd into no other".

causal origins of a passion, the passion produced is not a copy of its cause in the way an idea is a copy of the impression from which it is derived.¹⁴ Part of what it is to be an impression is to be an original, that from which copies are made. Passions, like all impressions, are not copies of anything else.

This still leaves us with the problem of the directedness of anger. Typically, I am angry at someone, and not just possessed by blind rage. But Hume has no problem in accounting for this. Many of the passions, such as pride and humility, love and hatred, take the self or another as their object. But a passion such as hatred doesn't itself represent another; rather it is associated with an idea of another, and Hume has no problem with ideas representing, as we have seen. Hume explicitly likens anger to hatred in always taking another as its object: " 'Tis the same case with hatred. We may be mortified by our own faults and follies; but never feel any anger or hatred, except from the injuries of others." (*T* 329-330)¹⁵

PART THREE: BELIEFS, REASON AND MOTIVATION

In 1.3.10, "Of the influences of belief," Hume introduces the subjects of motivation to action and the role that belief might play in motivation. We should remember that beliefs are the product of causal reasoning and, to put the matter crudely,

¹⁴ That Hume is using "original" in importantly different ways here was pointed out to us by Ashley McDowell. Also see the discussion of this difference in Rachel Cohon (1994), pp. 188-9.

¹⁵ Though Hume seems to speak of hatred and anger here and elsewhere as being virtually the same passion, it turns out later that anger is a separate passion that can be caused by hatred, in just the same way that benevolence is a separate passion that can be caused by love (*T* 368).

are distinguished from ideas of the imagination, or merely conceived ideas, by their greater force and vivacity.¹⁶ Hume says "There is implanted in the human mind a perception of pain and pleasure, as the chief spring and moving principle of all its actions. But pain and pleasure have two ways of making their appearance in the mind; of which the one has effects very different from the other." (T 118) That is to say, pain and pleasure may occur either as impressions or as ideas. The former "always actuate the soul, and that in the highest degree." But ideas have a variable effect. By and large, it is only beliefs "which produce in a lesser degree the same effect with those impressions, which are immediately present to the senses and perception" (T 119). So "[t]he effect, then, of belief is to raise up a simple idea to an equality with our impressions, and bestow on it a like influence on the passions. This effect it can only have by making an idea approach an impression in force and vivacity." (T 119) Impressions, especially impressions of pain and pleasure, influence our actions by virtue of their force and vivacity. By and large ideas, having much less force and vivacity, do not. Beliefs, being ideas with more force and vivacity, approximate impressions in their motivational strength. But beliefs are, in part, the conclusions of causal rea-

¹⁶ It is worth noting that Hume explicitly says that "I shall here anticipate a little what wou'd more properly fall under consideration afterwards, when we come to treat of passions and the sense of beauty." (T 118) Considering this explicit forward reference, and the fact that Books 1 and 2 were published together, it is highly unlikely that there is any serious conflict between what Hume says in this section and the first presentation of The Representation Argument. This does not rule out a subtle inconsistency, or a slightly less subtle inconsistency with the later formulation in the separately published Book 3, which we know was rewritten extensively. But it makes the apparent glaring inconsistency very unlikely.

soning. So the results of reasoning can have motivational force. At this stage, we make only these two observations about this point. First, Hume seems to be committed only to the limited claim that beliefs about *pain* or *pleasure* may be the conclusion of reasoning and may motivate. He is silent on the prospect of other beliefs motivating. Second, the point about beliefs' influence on passions or actions is not developed here by Hume. Indeed, he spends the bulk of this section discussing the effects of beliefs on the imagination.

We must now see how these remarks about the influence of belief can be made consistent with Hume's famous thesis of the motivational impotence of reason. The key to the explanation lies in the Representation Argument. Here is its structure:

- (1) Passions have no representative quality.
- (2) Only what represents real relations and matters of fact, and so can agree or disagree with them (*T* 458), can be contrary or conformable to reason.
- (3) Therefore passions cannot be contrary or conformable to reason.

We have seen that the controversial first premise is really just a consequence of passions being impressions rather than ideas. If we grant Hume the second premise, the conclusion follows: passions (and volitions and actions) cannot be contrary to reason or conformable to it.¹⁷ Passions and volitions

¹⁷ One may not wish to grant the second premise, of course; it has problems of its own, not addressed here. It restricts reason and its products to that which can represent. Hume may accept it because of considerations about what all reasoning processes or their products can have in common. Demonstration and probable inference are known activities of reason. There could be others; but it is hard to

cannot be contrary to reason because they are impressions and hence do not represent. Actions, whether or not they are impressions, are at any rate not ideas, so the same applies.

It is hard to see straight off how an argument proving that passions and actions cannot represent something else by resembling it, and in this sense cannot conform to reason, is supposed to show that reason alone cannot motivate actions. Yet this is clearly what Hume intends it to show. He introduces the Representation Argument to prove that "reason alone can never be a motive to any action of the will" (*T* 413). In 3.1.1 he says the Argument proves "that reason is perfectly inert, and can never either prevent or produce any action or affection" (*T* 458). And he uses this to show, both directly and indirectly, that moral distinctions are not derived from reason.

The famous indirect argument that moral distinctions are not derived from reason is formulated several times; at its first appearance it looks like this:

- (1) "... morals ... have an influence on the actions and affections..."
- (2) "... reason alone ... can never have any such influence."
- (3) "... it follows, that [morals] cannot be deriv'd from reason" (*T* 457).

On the next page the premise about the inertia of reason reads this way: "... reason can never immediately prevent or produce any action by contradicting or approving of it ..." This premise that reason alone cannot produce action, which we will call the Inertia

imagine that a non-representational mental process could have enough in common with these known activities of reason to be classed as reasoning, or its products as products of reason.

Thesis, seems to be a (negative) causal claim. The conclusion of the Representation Argument, from which it is supposedly derived, is that passions and actions are not accurate or inaccurate copies of any other reality which they purport to represent, which does not seem to be at all about what reason alone can cause. Hume clearly thinks he can move from the one claim to the other in a single step. How?

We should note that the Inertia Thesis does not say that beliefs alone cannot produce passions or actions, but rather that *reason* alone cannot. This may seem to be compatible with the motivational efficacy attributed to beliefs about likely pleasures and pains in 1.3.10. But such beliefs are conclusions of causal reasoning. So if these claims are to be compatible, we must be able to explain the following apparent conflict. Suppose that as the result of a bit of causal reasoning I believe that driving under the influence may well cause me pain, and this belief gives rise to an aversion which moves me to refuse the next drink. This causal sequence is an instance of belief motivating action in just the way described in 1.3.10. But it also seems to be an instance of causal reason alone motivating action, and so just the sort of thing ruled out by the Inertia Thesis.¹⁸ True, the passion of aversion intervenes in the causal sequence, but this is no help; for in most versions of the Inertia Thesis Hume says that reason alone cannot produce passions or volitions either (*T* 457-8). So we must also explain why this causal sequence, a sequence countenanced by Hume, does not

¹⁸ Baier renders 1.3.10 consistent with Books 2 and 3 by dismissing the Representation Argument (Baier (1991), p. 160). Michael Gill tries to demonstrate consistency by distinguishing two types of belief. See his "Reason, Belief and the Motivating Influences of the Will," read at the Hume Society Conference, Ottawa, 1993.

count as a case of reason alone producing passion, in contradiction to the Inertia Thesis.¹⁹

One enticing way to explain both how the Representation Argument supports the Inertia Thesis and how the Inertia Thesis is compatible with "Of the influence of belief" is to deny

¹⁹ This seems to be why Hume is read by so many as denying that belief can motivate without the help of some independent passion, one not caused by the belief (see, for example, J. L. Mackie (1980), p. 47, p. 52; Francis Snare (1991), p. 47, pp. 84-5). They do so not from neglect of 1.3.10, but from seeing no way to make what Hume apparently says there consistent with 3.1.1. It seems that in order to say that reason alone cannot produce passions or actions, Hume needs to say that beliefs alone cannot. The solution could be to discount some of the language of 1.3.10 and read 2.3.3 with appropriate emphasis, so that where Hume seems to say that a causal belief about the sources of pleasure or pain creates a new desire or aversion, what he means is that we have a constant general desire for pleasure and aversion to pain which is given a specific object by the belief that we can get or avoid these in certain ways. Another solution in the same spirit would be that when the idea of pain or pleasure from a certain source is enlivened to become a belief forceful enough to cause desire or aversion, it is no longer a mere belief but an actual mild feeling of pleasure or pain. Either way, a belief alone, even one about the sources of pleasure and pain, does not produce passion or action. But Hume does not actually say either of these things, and what he does say is that beliefs about the probable sources of pleasure and pain themselves cause volition and action. (See, e.g., *T* 414: "... the impulse arises not from reason, but is only directed by it. 'Tis from the prospect of pain or pleasure that the aversion or propensity arises towards any object ...") The intent here, then, is not to pursue either of these readings, but to take Hume at his word that mere belief about the likely future sources of pain and pleasure does cause passion and action, and to try to square this with the Inertia Thesis of 2.3.3 and 3.1.1. Hume seems to think that an impulse which arises from the prospect of pain or pleasure, a probabilistic belief, is not one that arises from reason alone.

that the thesis that reason is inert is a causal claim, as it appears to be. Since, as the Representation Argument shows, the products of reason are ideas which represent their originals, perhaps what we are to conclude from the Representation Argument is that the only products of reason are *conclusions* of reason, the outcomes of demonstrative or causal inferences. Thus the claim that reason alone cannot produce action is an ontological – almost a logical – thesis, rather than a causal one: passions, volitions, and actions cannot be the conclusions of bits of reasoning, because they are of the wrong ontological category, “realities” rather than representations that can be true or false. The Inertia Thesis in the indirect argument should then be understood to say that passions and actions cannot be entailed by premises or derived by inference, or (more broadly) that they cannot be produced by the recognition that they would be accurate representations.²⁰ If being produced by reason alone is being produced as a conclusion, then of course the causal sequence from causal inference to belief about the danger of drunk driving to aversion and thence to refusing the drink does not count as production of action by reason alone. Reason’s work is done once the belief is formed. The next step (from belief to passion) is not a piece

²⁰ There is some historical evidence that this was Hume’s intention. Hutcheson argued that good and evil in actions cannot be identical with reasonableness and unreasonableness in them on these grounds: the reasonable is the true and the unreasonable is the false; there are as many true propositions about evil as about good actions, and as many false ones as well; so good actions are no more true than evil ones, and hence no more reasonable (Francis Hutcheson (1971), p. 120). It is easy to imagine Hume taking over the beginning of this argument and then making the following move: actions are not the kinds of things to be true or false, therefore not the kinds of things to be reasonable or unreasonable, so the rationalist position is absurd.

of reasoning but mere causation, so this is not production by reason alone.

However, if what Hume means by the Inertia Thesis is that passions and actions cannot be conclusions of inferences, he equivocates in the indirect argument. Recall what it says:

- (1) Reason alone cannot produce passions or actions.
- (2) (Judgments of) moral merit and demerit can produce passions and actions.

Therefore, moral distinctions are not the offspring of reason alone. (T 458)

If we interpret 'reason alone' as we have so far, the first premise says:

- (1') Reason alone cannot produce passions or actions *as conclusions*.

But of course, the second premise does not say that moral distinctions *can* produce passions and actions as conclusions. (*This* is not how moral distinctions produce them, and in any case they cannot *be* conclusions, since they are original existences.) The productive influence of morality on passions and actions is merely causal. That is the equivocation.²¹

The danger of equivocation arises because the indirect argument apparently depends for its validity upon some sort of transitivity, some principle of the form "If A alone produces B,

²¹ For an account that attributes just this equivocation to Hume, see Rachel Cohon (1988).

and *B* (alone?) produces *C*, then *A* alone produces *C*." For the argument has the following structure:

- (1) *A* alone cannot produce *C*.
- (2) *B* produces *C*.
- (3) Therefore *A* is not the source of *B*.²²

These premises entail the conclusion only if the relation "produces" or "is the source of" is transitive. But with the equivocation, transitivity is lost.

So, simply reading the Inertia Thesis as saying that only representations can be the conclusions of bits of reasoning (a corollary of Premise 2 of the Representation Argument) makes the indirect argument invalid. Perhaps Hume does equivocate in this way. However, we will propose a reasonable alternative reading of the indirect argument which does better for Hume.

PART FOUR: REASON AS A KIND OF CAUSE, AND REASON ALONE

Let us return to the plausible supposition that the Inertia Thesis is really a causal thesis. This leaves us with at least two problems to solve. First, how is such a causal claim supported by the Representation Argument? Second, the causal

²² This is a schematic paraphrase of the following passage: "... as reason can never immediately prevent or produce any action by contradicting or approving of it, it cannot be the source of the distinction betwixt moral good and evil, which are found to have that influence." (*T* 458) Hume states what we call the indirect argument four or five times in "Moral distinctions not deriv'd from reason"; this is the third time. The exact wording varies slightly in the different repetitions, but in every formulation, Hume sounds as if he means to use the same verb in the two premises.

reading of the indirect argument presupposes the principle of causal transitivity. This principle apparently entails that if reason alone causes the cause of an action, this counts as reason alone causing action. In our case of refusing the drink, since causal reasoning produces the belief about pain, which causes aversion, which causes action, in light of the principle of transitivity it looks as if reason alone causes action. But then the motivational efficacy of beliefs in the prospect of pleasure or pain, as described in 1.3.10, is incompatible with the Inertia Thesis, that reason alone cannot cause action. For the sole difference cited in 1.3.10 between ideas of pleasure and pain that motivate action and those that do not is that the former are beliefs; there is no indication that they are aided in their motivational efficacy by any causally-independent passion or state.²³

This leads us to a third challenge. Merely avoiding the equivocation by invoking causation in both premises is not enough to insure that the argument is valid. Since the argument depends for its validity on the tacit assumption of the principle of the transitivity of causation, it succeeds only in cases where that principle holds. There are cases where the principle holds and cases where it does not. For example, in this argument, the transitivity of causation seems to hold and complete the inference:

- (1) Determination alone cannot make me rich.
- (2) Good luck alone can.
- (3) Therefore, determination alone cannot produce good luck.

²³ If beliefs about the prospect of pleasure or pain cause passions or actions only under particular conditions, then, it seems, there are hidden causes at work in those circumstances, and the beliefs are not causally sufficient for the passions or actions.

It does seem that if determination alone could produce good luck, then in a certain sense, determination would be sufficient to make me rich. But the following argument, seemingly of the same form, fails:

- (1) Being indicted, alone, cannot give Jones a headache.
- (2) Tension, alone, can give Jones a headache.
- (3) Therefore, being indicted, alone, cannot cause tension.

The premises are at least plausible, but the conclusion seems nonetheless false, and a *non sequitur*. This is because we do not naturally insert the principle of transitivity as a tacit premise, and indeed, that principle appears false here. We are not inclined to say that if being indicted alone could cause tension and tension alone could cause headache (in Jones), then being indicted alone would be sufficient to make Jones' head hurt. An interpretation that validates the indirect argument must construe the principle of transitivity in such a way that it is true, presumably by narrowing it so as to exclude this type of case.²⁴

To deal with these problems, we have to reconsider what "reason alone" means. It has been usual to interpret this as "beliefs, without passions or sentiments." This is incompatible

²⁴ Another way to provide a valid reading of the indirect argument is to interpret its conclusion as a noncausal claim and construe the inference not as passing through causal transitivity but as having some other logical form. Reading the argument as not requiring the principle of transitivity is one way to avoid the problems introduced by that principle. For such an interpretation, see Rachel Cohon (1997), pp. 251-266. Here we pursue the causal interpretation of the conclusion.

with 1.3.10. When reason is contrasted with the passions, or with the moral sentiments, Hume is contrasting operations of the understanding, which deal only with ideas, with those operations that also concern the sentiments. So reason alone, we suggest, is reasoning *considered apart from any passions and any feelings of pleasure or pain*. On this view, reason is a kind of cause, to be sure, but one whose effect is truths (T 180), and falsehoods when other causes intervene, but in any case ideas rather than "realities."²⁵ Reason alone, so understood, can give rise only to representations.

The Representation Argument is indeed offered as support for the Inertia Thesis in at least two places (T 415, T 458). The Representation Argument reasserts the claim about the nonrepresentational nature of the passions, characterizing them as original existences. But reason is concerned with the relations of ideas, relations that hold either between ideas themselves, or between an idea and some "*real existence and matter of fact*" (T 458). We can reason *about* passions, just the way we can reason about any real existence. But we can only reason *with* ideas, that is, with entities capable of representing other realities. "Now 'tis evident that our passions, volitions, and actions, are not susceptible of any such agreement or disagreement; being original facts and realities, compleat in themselves, and implying no reference to other passions, volitions, and actions" (T 458). But, the Representation Argument continues, from this claim about passions, representations, and realities, it immediately follows that passions cannot be conformable to or opposed by reason: "'Tis impossible, therefore, they can be pronounced either true or false, and be ei-

²⁵ "Our reason must be consider'd as a kind of cause, of which truth is the natural effect ..." T 1.4.1.

ther contrary or conformable to reason" (*T* 458).²⁶ Reason produces only ideas or representations. But passions, volitions and actions are "real existences" or "original facts and realities, compleat in themselves." Hence they can never be the outcome of reason. Nor can they be conformable to or opposed by any outcome of reason.²⁷

A creature with reason alone, in the sense we are suggesting Hume intended, would be one who had Humean reason but no passions or feelings of pleasure or pain.²⁸ Such a being would, of course, come to have some beliefs. But such a being could not have ideas of the passions, nor of pleasure or pain, since he could not experience the originals. Consequently he could not form any beliefs about them, even if his causal reasoning were perfect. Understood in this way, reason *alone* does not produce any beliefs about the prospects of pain or pleasure either. Reason cannot produce impressions, nor can it produce any new ideas (*T* 157). Thus, reason alone cannot produce the one kind of belief that on Hume's account is causally linked with passion and action.

Of course, in a being also possessed of feelings of pleasure and pain, and of the capacity to desire the one and shun the other, reason plays an important role in the production of beliefs about the sources of pleasures and pains, and these be-

²⁶ Compare *T* 415: "'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason; since this contradiction consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent."

²⁷ On the thesis that reason can cause only what is conformable to or contrary to it, see Cohon (1997).

²⁸ For more details about such a Humean pure reasoner, see David Owen (1994).

liefs, in turn, cause passions and action. But such beliefs will not be the products of reason *alone*, even in that being, for in forming such beliefs it would also need to make use of ideas not available to reason alone. This makes the inertia of reason alone compatible with the motivational efficacy of certain beliefs. Beliefs of the motivating kinds are ones that reason alone cannot produce.

The indirect argument against moral rationalism then must be understood to say this:

- (1) Reason alone, given what it is, cannot give rise to any actions, nor to any beliefs of the kind that in fact cause action.
- (2) Moral judgments do cause actions.
- (3) Therefore reason alone does not give rise to moral judgments.

So, even if moral judgments should turn out to be beliefs, they will not be beliefs of the kind that are produced by reason alone. To form such beliefs one needs to make use of passions, or feelings of pleasure or pain, or both.

The argument nonetheless still relies on the principle of transitivity, which is not generally valid. But the principle is valid on our understanding of "reason alone." Recall the argument about determination and wealth:

- (1) Determination alone cannot make me rich.
- (2) Good luck alone can.
- (3) Therefore, determination alone cannot produce good luck.

“X alone can produce Y” seems to mean here that X is sufficient to produce Y without the contribution of any independent causal factor.²⁹ If, in this sense, X alone (by itself) can cause Y, and Y by itself can cause Z, then it seems that X by itself can cause Z, even though it does so by way of Y, for Y is entirely within X’s causal control. Y is not an independent contributor to the process. Thus, if determination by itself *could* produce good luck for me (if Mark Twain’s aphorism, “The harder I work, the luckier I get,” were a literally true causal claim), and good luck by itself can make me rich, then determination alone could make me rich, in this sense of “alone.” Since determination is not sufficient to make me wealthy without the contribution of some other causal factor not caused by my determination, then it follows that, sadly, good luck is not under the causal control of determination. This is an analog to the indirect argument as we interpret it. If reason alone could produce moral distinctions, which we know can, alone, produce passions and actions, then reason alone would be capable of producing passions and actions. Since reason cannot produce passions or actions without the additional contribution of something not caused by reason alone, it follows that moral distinctions are not under the causal control of reason alone.

In the argument about tension and headaches, however, this is not the sense of “alone” in virtue of which we take the first premise to be true. Recall that argument:

²⁹ Unfortunately, one can easily slip into reading ‘X alone can produce Y’ as if it said “only X (and nothing else) can produce Y.” Of course, this is not what we mean by ‘alone’; as in Hume’s argument about reason and moral judgments, ‘X alone’ here, and in the argument about determination and luck, means “X by itself.”

- (1) Being indicted, alone, cannot give someone a headache.
- (2) Tension, alone, can give someone a headache.
- (3) Therefore, being indicted, alone, cannot cause tension.

Were we to assert the first premise, we would presumably mean that being indicted does not cause a headache without an intermediate step: there are other things happening between the indictment and the throbbing of the head, such as tension in the neck or a reduction or increase in the blood flow. However, this is different from the sense of "alone" that made the former arguments valid. The sense used here does not support the principle of transitivity. From the facts that *X* causes *Y* without intermediary, and *Y* causes *Z* without intermediary, it does not, of course, follow that *X* causes *Z* without intermediary. We propose that what Hume means by "reason alone" in the indirect argument is thus "reason understood apart from pleasure, pain, and the passions, and without independent causal contribution from anything which reason cannot cause;" but not "reason directly, not operating through any intermediate products."³⁰

³⁰ We are grateful to Gerald Dworkin, who directed our attention to the need to make the principle of transitivity more precise by proposing this counterexample:

Reason alone cannot cause children's actions.

Parental orders alone can cause children's actions.

Therefore, parental orders are not the products of reason alone.

This is clearly invalid. But the sense of 'reason alone' that makes the conclusion false is not Hume's sense. Indeed, Hume would say that reason alone in his special sense cannot produce parental orders any more than it can produce any other action. If when we read this

PART FIVE: CONCLUSION

In this paper we presented the Representation Argument as a consequence of a more general thesis about impressions, ideas and representation. Our interpretation of the indirect argument against moral rationalism treats the Inertia Premise as causal, thus avoiding the equivocation problem. It also sees that premise as following from the Representation Argument. This requires a new reading of “reason alone”: reason alone is reason functioning in isolation from any passions, sentiments, or feelings of pleasure and pain, either in the form of impressions or ideas, and apart from anything that reason itself (so understood) cannot cause. This insures that the indirect argument relies only on the true version of the principle of transitivity. Our reading of the Inertia Premise is entirely compatible with 1.3.10: beliefs about pleasure and pain can motivate, but such beliefs cannot enter into arguments of reason alone. One consequence of our interpretation, which we mention here only in passing, is that there is nothing in the indirect argument to support a non-cognitivist reading of Hume. Moral judgments may well be beliefs for Hume. It is just that they are not the sort of beliefs that can be reached by reason alone.

argument we think it can, this is not necessarily because we have a certain philosophical position on motivation by reason. The contrast between parental orders, which can be products of reason “alone” in a certain sense, and children’s actions, which cannot, is that reason can cause children’s actions only through parental orders, whereas parental orders can be generated directly by reason without an intervening step. As for whether reason can generate parental orders without an additional, independent causal factor operating simultaneously with reason, the argument does not address that issue. But that is actually the issue with which Hume is concerned.

REFERENCES

- BAIER, A. (1991). *A Progress of Sentiments: Reflections on Hume's "Treatise"*. (Cambridge, Mass., Harvard University Press).
- COHON, R. (1988). Hume and Humeanism in Ethics. *Pacific Philosophical Quarterly*, vol. 69, n° 2 (June).
- . (1994). On an Unorthodox Account of Hume's Moral Psychology. *Hume Studies*, vol. XX, n° 2, 179-194.
- . (1997). Is Hume a Noncognitivist in the Motivation Argument? *Philosophical Studies*, vol. 85, 251-266.
- GILL, M. (unpublished). Reason, Belief, and the Motivating Influences of the Will, (read at the Hume Society Conference, Ottawa 1993).
- HUTCHESON, F. (1971). *Illustrations on the Moral Sense*, ed. Bernard Peach (Cambridge, Mass., Harvard University Press).
- LOCKE, J. (1975). *An Essay Concerning Human Understanding*, ed. P.H. Nidditch (Oxford, Clarendon Press).
- MACKIE, J.L. (1980). *Hume's Moral Theory* (London, Routledge).
- MILLGRAM, E. (1995). Was Hume a Humean?, *Hume Studies*, Vol. XXI n° 1, 75-94.

OWEN, D. (1994). Reason, Reflections and Reductios, *Hume Studies* Vol. XX n° 2, 195-210.

SNARE, F. (1991). *Morals, Motivation, and Convention* (Cambridge, Cambridge University Press).