# Hume’s (*Ad Hoc*?*)* Appeal to the Calm Passions

*Hsueh Qu*

*National University of Singapore*

## Conativism: Relation of Ideas or Matter of Fact?

Hume famously champions conativism, that is, the position that the passions do the end-setting motivational work, while reason[[1]](#footnote-1) plays only an instrumental role by revealing how to achieve the goals set by our passions.[[2]](#footnote-2) He states that ‘reason alone can never produce any action, or give rise to volition’ (THN 2.3.3.4).[[3]](#footnote-3) ‘Reason is, and ought only to be the slave of the passions’ (*ibid*.) and only serves to discover the connection between causes and effects, facilitating the achievement of the goals set by our passions; in short, reason plays only an instrumental role in motivation, while the passions play the end-setting role. I will henceforth use ‘conativism’ and its cognates to indicate only the thesis that the passions play the end-setting motivational role; the details of the instrumental role are tangential to my purposes here.

Conativism is subject to apparent counterexamples in which the phenomenological evidence supports the rationalist claim that reason performs the end-setting role in motivation. Imagine a student who reluctantly decides to revise for an upcoming test instead of attending a party – to her it *feels* as if she has a passion for the latter but none for the former, with reason (successfully) telling her to revise; the phenomenology therefore suggests that reason successfully motivates in an end-setting role. In response to such purported counterexamples, Hume defends his conativism by appealing to the calm passions:

... there are certain calm desires and tendencies, which, ’tho they be real passions, produce little emotion in the mind, and are more known by their effects than by the immediate feeling or sensation… When any of these passions are calm, and cause no disorder in the soul, they are very readily taken for the determinations of reason, and are suppos’d to proceed from the same faculty, with that, which judges of truth and falshood. (THN 2.3.3.8)

Hume argues that in these apparent counterexamples, there are one or more passions that perform the end-setting role.[[4]](#footnote-4) Because these passions are ‘calm’ and produce little discernible emotion, however, we mistakenly think reason to be doing the motivational heavy-lifting.[[5]](#footnote-5),[[6]](#footnote-6) Hume goes on to explain that a calm passion arises when a passion becomes ‘a settled principle of action’; ‘repeated custom and its own force have made every thing yield to it’, thus removing the ‘opposition and emotion’ that would typically accompany it (THN 2.3.4.1). He also notes that passions tend to be calm when their objects are distant: ‘The same good, when near, will cause a violent passion, which, when remote, produces only a calm one’ (*ibid*.). Hume thus provides an explanation of how the calm passions would fit into his psychological account.

Despite Hume’s efforts, Cohon (2008, p.17) notes that this appeal to the calm passions seems ‘inadequate’, so long as it is read empirically; similarly, Stroud (1977, p.167) argues that ‘Hume does not prove, or even make very plausible, the contention that a feeling or passion must always be present for action to occur’, which indicates his dissatisfaction for Hume’s empirical case for the calm passions. If the calm passions lack adequate empirical support, then it seems that appealing to them in order to rescue conativism seems an *ad hoc* move.

To circumvent this worry, Stroud (1977, pp.168-9) departs from Hume in providing a functionalist[[7]](#footnote-7) account of desires.[[8]](#footnote-8) Hume’s account is seemingly problematic as an empirical one because its appeal to the calm passions lacks adequate empirical support, and thus seems *ad hoc*. Dispositional accounts such as Stroud’s do not ground the conativist thesis empirically, as desires are *defined* as being the end-setting component of motivation, thus removing the need for empirical confirmation: by definition, any case of motivation *must* involve a passion in an end-setting capacity. In essence, such accounts differ from Hume’s with respect to the status of the conativist thesis: in the terminology of ‘Hume’s Fork’ (EHU 4.1), Hume believes this thesis to be a matter of fact (that is, empirical), while Stroud construes it as a relation of ideas (that is, conceptual). Viewed as a matter of fact, Hume’s appeal to the calm passions can seem question-begging given the lack of empirical support for this claim. Stroud circumvents this problem by treating the conativist thesis *conceptually* – a passion is *whatever* is the end-setter in motivation, avoiding any such problem.

Similarly, the ‘direction of fit’ reading also views the conativist thesis as a relation of ideas. Bricke (1996) adopts Michael Smith’s (1994) Humean account whereby beliefs and desires differ in having different directions of fit[[9]](#footnote-9) – beliefs aim to conform to the world and therefore have a mind-to-world direction of fit, while desires aim to make the world conform to them and therefore have a world-to-mind direction of fit.[[10]](#footnote-10) Crucially, Bricke’s functionalist account diverges from a strictly phenomenological conception of the passions.

In what follows, I will argue that conceptual accounts of Hume’s conativism fail both as interpretations of Hume and as theories in their own right, the cure being far worse than the disease. I then argue that an empirical reading can in fact be theoretically justified against objections of *ad hockery*, both by Hume’s rules by which to judge of causes and effects (THN 1.3.15), and by the framework of the theoretical virtues.

## Trivial Pursuits and Other Problems

One initial worry for conceptual readings of Hume’s conativist thesis is that in trying to evade the frying pan of *ad hockery* on Hume’s behalf, they cast him instead into the fires of triviality. The conceptual reading of the conativist argument designates the passions (construed dispositionally) as the end-setting cause of motivation. However, citing a bare disposition to an effect as a cause of this same effect is no explanation at all, but merely redescribes the explanandum. Stroud’s formulation of his position clearly highlights this emptiness:

It might well be that to have a desire for or propensity towards E is simply to be in a state such that when you come to believe that a certain action will lead to E you are moved to perform that action... *And being in some such dispositional state might be all that having a certain desire or propensity consists in.* *It need not be an additional mental item that itself produces the action*. (Stroud 1977, pp.167-8, emphasis mine)

If a desire is wholly constituted by its dispositional property of ‘being the end-setting cause of motivation’, it trivially follows that desires play the end-setting role in motivation. Similarly, the ‘direction of fit’ reading of Hume’s conativist thesis characterises desires as whatever aims to make the world conform to them; that is, it characterises desires as exactly those things that perform the end-setting role in motivation! Again, it becomes trivially true that any case of motivation will involve a passion performing this end-setting role. Bricke’s discussion, like Stroud’s, highlights this shallowness of his account:

Were they genuinely beliefs, the cognitivist’s explicitly evaluative beliefs could not play the major role in reasons for action for they would have the mind-to-world direction of fit. On the other hand, if they can play the major role in action explanations – and it seems clear that they can – they cannot be beliefs, for, playing that major role, they must have the world-to-mind direction of fit. (Bricke 1996, p.28)

Here Bricke argues that evaluative beliefs *must be* desires simply because they play the end-setting (or major) role in motivation – in short, whatever plays the end-setting role in motivation must be a desire, simply *in virtue of what it is to be a desire*. This smacks of triviality.

Hume, needless to say, is no fan of trivial explanations. He is trenchantly critical of the ancient philosophers’ usage of mere dispositions to ‘explain’ unknown phenomena:

By this means, these philosophers set themselves at ease, and arrive at last, by an illusion, at the same indifference, which the people attain by their stupidity, and true philosophers by their moderate scepticism. They need only say, that any phaenomenon, which puzzles them, arises from a faculty or an occult quality, and there is an end of all dispute and enquiry upon the matter. (THN 1.4.3.10)

Indeed, Hume’s conativist argument cannot be trivial, because he intends it to do substantial philosophical lifting, refuting his rationalist opponents who believe reason alone sufficient to motivate action:

Nothing is more usual in philosophy, and even in common life, than to talk of the combat of passion and reason, to give the preference to reason, and assert that men are only so far virtuous as they conform themselves to its dictates... In order to shew the fallacy of all this philosophy, I shall endeavour to prove *first*, that reason alone can never be a motive to any action of the will; and *secondly*, that it can never oppose passion in the direction of the will. (THN 2.3.3.1)

Interpreted conceptually, Hume’s argument for conativism spectacularly fails to accomplish his stated goal of refuting (or even engaging with) his rationalist opponents; for example, the rationalists could simply accept such a definition of ‘passion’ as whatever plays the end-setting role in motivation, but claim that reason is a passion.

Bricke unsurprisingly denies such an accusation of triviality:

Humean desires… must be characterizable in a suitably independent way. But they are so characterizable. They are ineliminable elements, psychologically primitive elements, in an explanatory theory of human behaviour. (Bricke 1996, p.31)

It is unclear whether this gloss genuinely characterises Humean desires independently from describing them as having a world-to-mind direction of fit, given that their primary role in ‘an explanatory theory of human behaviour’ seems to be that they play the ‘major role in action explanations’, which is inextricably tied to their world-to-mind direction of fit, as Bricke’s quotation above (p.28) makes clear. Of course, since Bricke seeks to characterise desire functionally, he cannot independently characterise them by appealing to their qualitative character. Nevertheless, even if his account does indeed provide an independent functional characterisation for desires, the thesis would then violate Hume’s principle of the metaphysical contingency of causation (henceforth CC): ‘*a priori*, any thing may produce any thing’ (THN 1.4.5.30). Conceptual readings of the conativist thesis face the following dilemma: citing a bare disposition as a cause renders the thesis empty and lacking in explanatory value, but fleshing out the disposition by attributing substantive properties to it violates CC. I take this to be a philosophical problem as well as an interpretive one, given CC’s plausibility and widespread acceptance.[[11]](#footnote-11)

This dilemma can be highlighted by the following example. The conceptual truth that ‘egg-breakers break eggs’ (where ‘egg-breakers’ are defined such that all and only egg-breakers break eggs) does not violate CC, because it does not make any substantive claim about the cause of broken eggs, merely assigning it a name – the concept of an egg-breaker is bare other than its functional characterisation. The statement therefore does not *a priori* rule anything out as the cause of broken eggs, which renders it consistent with CC, but at the cost of triviality. However, if we avoid emptiness by attributing substantive properties to the notion of an ‘egg-breaker’, we would instead violate CC. Similarly, the conceptual reading of the conativist thesis faces an identical dilemma: either it defines desires as pure dispositions to perform the end-setting motivational role, in which case the thesis would be trivial; or it defines desires as more, in which case it would violate CC.[[12]](#footnote-12) And indeed, Hume sees his account of the passions as being empirically substantive, impaling him on the second horn: the passions are occurrent mental states with qualitative characters, for example.

Note that I am not attacking functional characterisations *per se*, merely *a priori* ones; I think empirical functional characterisations to be unproblematic. For example, the Higgs-Boson particle is defined functionally as something like ‘the particle that grants elementary particles their mass’, yet its postulation is highly respectable. Such a functional characterisation does not render the thesis *M* that ‘the Higgs-Boson grants elementary particles mass’ trivially true (on the assumption that mass does exist), because the Higgs-Boson is also characterised as *a sub-atomic particle* *that is* *a vibration of the Higgs field under high energy conditions*, for example.[[13]](#footnote-13) The Higgs-Boson is not merely characterised as ‘whatever grants mass’; if it turned out instead that mass was granted by tiny fairies weighing down particles, physicists would hardly claim that they had discovered that Higgs-Bosons were fairies! In short, *M* is not an *a priori* truth, since we have to look into the world to verify it, examining if there is anything out there that grants mass and matches our description of a Higgs-Boson; *M* therefore does not violate the contingency of causation. Thus, *M* avoids the above dilemma.

For a functional construal of the conativist thesis to similarly avoid my dilemma, desires must be independently characterisable apart from their functional role, and the thesis must be empirically verified. But if these conditions are satisfied, then the conativist thesis has to be empirical rather than conceptual; we need to observe instances of motivation, examining the properties of the end-setting state and seeing if they match our (substantive) descriptions of desires.[[14]](#footnote-14) I take such an empirical functional characterisation of conativism to be unproblematic, but by itself, it does nothing to rescue Hume from the accusation of *ad hockery*. Examine the assiduous student: the conativist will claim that what is playing the end-setting role is a calm passion, while the rationalist will claim that it is a belief. To decide the issue, we must examine the properties of the end-setter to see which it is more likely to be (just as we examine the properties of the mass-granter and find it to fit our characterisation of a Higgs-Boson). Both theories agree in what they expect the functional role of this entity to be, so this cannot decide between the two accounts; thus we must look to phenomenology to decide the issue: does it *feel* like a desire or belief? And the answer is clearly the latter, leaving untouched the worry of *ad hockery*. An *a priori* functional characterisation solves the worry of *ad hockery*, but on pain of facing a dilemma between triviality and violating CC; an empirical functional characterisation avoids this worry, but does nothing to address the original objection. Therefore, functional characterisations by themselves do nothing to solve this problem.[[15]](#footnote-15)

## Rescuing an Empirical Account

### Making a Phenomenological Case for the Calm Passions

In judging the force of the objection of *ad hockery*, one crucial issue is whether the calm passions are sensible or not. Some commentators think Hume’s calm passions to be completely insensible: Stroud (1977, pp.164-5) claims that there is no independent way of discerning the calm passions besides their purported effects; Smith (1994, p.112) reads Hume as claiming that the calm passions ‘lack phenomenological content altogether’; Shaw (1989) believes the calm passions to be completely insensible, consequently arguing that they are unactualised dispositions to have certain desires were the individual to think about the relevant objects. The point is that Hume, as an empirically grounded scientist of man, only seems licensed to postulate sensible mental entities; if the calm passions lack any phenomenology, then Hume would be begging the question in postulating them. If these commentators are correct, then there would be no rescuing Hume from the objection of *ad hockery* – his appeal to the calm passions would be devoid of any empirical support and thus utterly question-begging. This would be especially troubling given Hume’s psychological framework – how can a passions exist independently of its being perceived in some manner?

I think Hume’s position not so dire: in appealing to the calm passions, Hume is not postulating something for which there is absolutely no phenomenological evidence. Hume clearly holds that we do on some level feel the calm passions. For one, he says they produce ‘little emotion’ (THN 2.3.3.8), and not no emotion. Likewise, Hume’s claim that the calm passions ‘cause no disorder in the soul’ (THN 2.3.3.8) merely indicates that they are not emotionally violent; a feeling of contentment could be said to ‘cause no disorder in the soul’, despite uncontroversially possessing phenomenal character. Similar remarks may be made about the claim that they produce ‘no sensible agitation’ (THN 2.3.4.1).

Indeed, that the calm passions have phenomenology can be verified by the fact that Hume takes the calm passions to be readily confused with the determinations of reason, precisely because they share a similar sensation:

 When any of these passions are calm, and cause no disorder in the soul, they are very readily taken for the determinations of reason, and are suppos’d to proceed from the same faculty, with that, which judges of truth and falshood. Their nature and principles have been suppos’d the same, because their sensations are not evidently different. (THN 2.3.3.8)

Clearly, the determinations of reason are not wholly insensible. Beliefs are characterised by their force and vivacity, which is a phenomenological notion: ‘An opinion, therefore, or belief may be most accurately defin’d, A lively idea related to or associated with a present impression’ (THN 1.3.7.5).[[16]](#footnote-16) Since beliefs have a distinct phenomenology, and the calm passions are confused with beliefs in virtue of sharing a similar phenomenology or sensation, it follows that Hume holds that the calm passions have phenomenology. Hume does say that reason ‘exerts itself without producing any sensible emotion’ (THN 2.3.3.8), but Hume only means that reason does not typically produce passions, rather than that it does not possess a phenomenology. The determinations of reason, like the calm passions, are perfectly sensible, but simply have a phenomenological valence that lacks violence. This reading can be confirmed by Hume’s noting that we have no need to invent unperceived passions, which clearly indicates that he believes that we do recognise the calm passions, even if only dimly:

Our passions are found by experience to have a mutual connexion with and dependence on each other; but on no occasion is it necessary to suppose, that they have existed and operated, when they were not perceiv’d, in order to preserve the same dependence and connexion, of which we have had experience. (THN 1.4.2.20)

Thus, it is plain that Hume does not believe the calm passions to exist unperceived.

Thus, the crucial issue in this debate is not whether Hume thinks the calm passions are sensible, because he clearly does. What is important is whether there is sufficient empirical grounds for taking ourselves to be motivated by calm passions rather than the determinations of reason. Certainly, we all feel *something* in cases such as the assiduous student, but is what we feel a calm passion or a determination of reason? If it seems more like a determination of reason than a calm passion, then Hume’s appeal to the calm passions can nevertheless be seen as *ad hoc*.

Nevertheless, note that the calm passions are not wholly indistinguishable from the actions of reason. When Hume claims that the two ‘are not immediately distinguishable to the feeling and perception’ (THN 2.3.3.8), this does not entail that they are completely indistinguishable, but rather that such a distinction is difficult and not immediate; similarly for his claim that ‘their sensations are not evidently different’ (ibid.), which is best read as claiming that their qualitative difference is not obvious, but nevertheless exists. Hume does not think this confusion between the calm passions and the determinations of reason to afflict everyone, but only those who lack a ‘strict philosophic eye’ (THN 2.3.3.8). Those with a less refined sense of introspection are liable to confuse the two, but those with more sophistication will be able to tell the difference.

A model for this can be found in Baillie (2000, p.95), who suggests that a refined observer would be able to discern the calm passions amidst the tempest of more violent emotions, a reading somewhat inspired by Hume’s adaptation of a tale from Cervantes’ Don Quixote in ‘Of a Standard of Taste’. Here Hume gives the example of the refined palate of Sancho Panza’s kinsmen, who discerned the subtle tastes of iron and leather in a hogshead of wine, which failed to be recognised by the less refined drinkers around them (EMPL 234-5). Just as an unrefined drinker might, in tasting a rich Barolo, mistake the a note of cherry for the taste of strawberry amidst the complexity of flavour, an unrefined feeler might mistake a calm passion for an action of reason. This phenomenon of greater and lesser sensitivity to our passions is described by Hume in his essay ‘Of the Delicacy of Taste and Passion’. Immerwahr (1991, p.9) argues that in this essay, ‘passion’ corresponds to the violent passions, while ‘taste’ corresponds to the calm passions. On this reading, Hume in this essay expresses approval of those who have a sensitivity to calm (but not violent) passions, presumably including those able to distinguish their calm passions from the actions of reason.

It should also be noted that Hume certainly thinks it possible for us to make second-order mistakes (such as misclassifying our perceptions) when reflecting on our perceptions:

... the operations of the mind... though most intimately present to us, yet, whenever they become the object of reflection, they seem involved in obscurity... (EHU 1.13)

But the finer sentiments of the mind, the operations of the understanding, the various agitations of the passions, though really in themselves distinct, easily escape us, when surveyed by reflection... (EHU 7.1).

On this reading, those who confuse the calm passions for the determinations of reason would be making a classificatory mistake along these lines – an understandable one, but a mistake nonetheless.

To sum up, intuitions seem split in cases of motivation involving the calm passions: some assiduous students, when tempted, will report the existence of a faint desire to stay home and study, while others might take this to be a belief or something similar. Hume maintains that the latter group are making a classificatory mistake, but is he justified in claiming this? For the sake of argument I will grant that plausibly, a significant proportion of people take themselves to be motivated by a belief or somesuch rather than a calm passion. In such a situation, Hume’s appeal to the calm passions, although not totally devoid of empirical support, would nevertheless be inadequately substantiated. This is problematic for an empirical construal of Hume’s conativist thesis: if one propounds an empirical thesis, one’s arguments for it had better have sufficient empirical grounding. In what follows, I proceed to make a theoretical case for justifying Hume’s appeal to the calm passions in light of this.

### Making a Theoretical case for the Calm Passions

Having addressed the phenomenological issue, I now argue that Hume’s appeal to the calm passions can be justified within such an empirical reading from the point of view of his methodology of systematisation. I provide two theoretical frameworks within which his appeal to the calm passions may be justified: first, by the admittedly modern framework of the theoretical virtues (of which there are hints in Hume),[[17]](#footnote-17) and secondly, by his own ‘rules by which to judge of causes and effects’ (THN 1.3.15). Note that such a theoretical evaluation is distinct from Hume’s arguments for his conativist thesis. Theoretical considerations come into play only after Hume’s arguments have made a case for the conativist thesis; the theoretical comparison of competing theses is only meaningful once the case has been staked for the theses and their respective theoretical credentials established.

### Systematisation

It is important to note the prominent role that systematisation – that is, the derivation of general principles from a multitude of observations – plays in Hume’s philosophical project.[[18]](#footnote-18) The crucial idea is that some general principles can be sufficiently well established that they are resistant to apparent counterexamples up to a point.

Systematisation is a recurring theme in Hume’s philosophy. In describing the *Treatise* in his anonymous *Abstract*, Hume emphasises the systematising nature of his project:

He proposes to anatomize human nature in a regular manner, and promises to draw no conclusions but where he is authorized by experience... This treatise therefore of human nature seems intended for a system of the sciences. (A 1-3)

Hume also notes that philosophy is simply this process of systematising our careful observations of our experience and reflections: ‘philosophical decisions are nothing but the reflections of common life, methodized and corrected’ (EHU 12.25). As Schliesser (2007) points out, for Hume, ‘experiment’ seems almost indistinguishable from systematic observations of our everyday experience. Demeter (2012, p.580) describes Hume’s project pithily: ‘Observation of human life is thus the point where inquiry and common life turn out to be continuous, the only difference being that the former is reflective, systematic and theory oriented whereas the other is unreflective, sporadic and practice-oriented.’ In short, Hume’s experimental method is clearly concerned with systematisation.

For Hume, systematisation involves the derivation of general rules from a multitude of observations.[[19]](#footnote-19) He explicitly describes his project in the *Treatise* as ‘being an attempt to introduce the experimental method of reasoning into moral subjects’, and describes this experimental method as consisting in ‘deducing general maxims from a comparison of particular instances’ (EPM 1.10). Similarly, in the first *Enquiry*, Hume identifies with the abstruse philosopher, whom he describes as deriving general principles from his observations of human nature:

...[the abstruse philosophers] regard human nature as a subject of speculation; and with a narrow scrutiny examine it, in order to find those principles, which regulate our understanding, excite our sentiments, and make us approve or blame any particular object, action, or behaviour... [P]roceeding from particular instances to general principles, they still push on their enquiries to principles more general, and rest not satisfied till they arrive at those original principles, by which, in every science, all human curiosity must be bounded. (EHU 1.2)[[20]](#footnote-20)

Likewise, ‘the utmost effort of human reason is, to reduce the principles, productive of natural phaenomena, to a greater simplicity, and to resolve the many particular effects into a few general causes, by means of reasonings from analogy, experience, and observation’ (EHU 4.12). Hazony (2014, p.162) stresses the centrality of the project of reducing to the fewest principles in Hume’s philosophy. Demeter (2012, pp.586-7) reads Hume’s experimental method as consisting in ‘a kind of *analysis* and *synthesis’* in the spirit of Newton: we derive principles from analysing phenomena, and we explain phenomena by synthesis, applying these principles to account for the phenomena. Hume’s reliance on analysis and synthesis is echoed by Hazony (2014, p.164), although he argues that Hume’s terminology and method ultimately more closely follows Boyle rather than Newton.

Hume’s emphasis on systematisation has two significant implications for his conativist thesis. First, it reveals its empirical nature: he is examining a multitude of observations about human motivation, and then deriving general principles from them. This is an unambiguously empirical undertaking founded on scientific observation, and his resulting principles must therefore be empirical generalisations rather than conceptual truths. Correspondingly, Hume’s arguments for this thesis are also empirical in nature: his dismissal of probable and demonstrative reasoning as end-setters in motivation (THN 2.3.3.2-3) rests on concrete examples (such as a merchant employing demonstrative reasoning only to guide his causal inferences) and introspection (in arguing that causal reasoning serves only to guide our passions). However, Hume’s argument that passions are ‘original existences’[[21]](#footnote-21) and therefore cannot oppose reason (THN 2.3.3.5) is sometimes seen as a conceptual one:[[22]](#footnote-22)

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. (THN 2.3.3.5)

I think that it would be a mistake to treat the above argument as conceptual in nature. The conclusion that Hume wishes to establish is that passions cannot ‘be oppos’d by’ reason; in short, reason cannot exert a countervailing motivational force to the passions. Importantly, this causal conclusion is the one that Hume needs if the argument is to support his conativism; a conclusion that (say) the passions are not truth-apt would not by itself bear on the topic of motivation, which THN 2.3.3 concerns. As previously mentioned, this opposition is a *causal* notion, and according to CC, *a priori* anything may cause anything. So, any argument for the claim that reason cannot oppose the passions would have to be empirical rather than conceptual in nature. Indeed, the argument rests on the observation that passions are original existences and are not copied from anything, which is plainly an empirical point. In examining our passions such as pride and humility, we find that the ideas that resemble them are copied from them rather than vice versa, as Hume establishes by empirical argument (introspection, citing the case of a blind man, etc.) in THN 1.1.1.8-9 in arguing for his Copy Principle. In short, Hume’s arguments for conativism are one and all empirical in nature.

The second crucial implication derives from the fact that a methodology of systematisation presupposes some theoretical criteria. A given set of data or intuitions may generate many competing sets of general principles, and we need some way to compare and choose between these different theories; systematisation is therefore useless without such criteria. Through such theoretical criteria, I will justify Hume’s appeal to the calm passions, construed empirically.

### The Theoretical Virtues

The theoretical virtues are certain properties that we think successful theories should possess; the best theory will be the one with the best balance of these virtues. I argue that Hume endorses three theoretical virtues in his philosophy: empirical confirmation, explanatory power, and simplicity.[[23]](#footnote-23) Empirical confirmation is the degree to which a theory is substantiated by the observed evidence,[[24]](#footnote-24) and explanatory power is how much the theory explains, and how little it leaves unexplained.[[25]](#footnote-25) Simplicity is (ironically!) a complicated virtue, and takes multiple forms.[[26]](#footnote-26) The driving idea behind simplicity being a theoretical virtue is that there are relatively few things (and/or types of things) in the world, and these things relate to each other in simple ways. Given that the world should contain relatively few things (and/or types of things), a theory should be *ontologically* simple; given that these things should relate in simple ways, a theory should be *syntactically* simple.[[27]](#footnote-27) A theory is ontologically simpler than another if it postulates fewer entities, or fewer kinds of entities (quantitative and qualitative parsimony respectively, in Lewis’ [1973a] terminology), and syntactically simpler if its basic principles are fewer and/or more concise, that is, if it describes entities as relating to each other in simple ways.[[28]](#footnote-28)

Hume takes empirical confirmation to be a crucial theoretical virtue. Boehm (2013, p.16) points out that ‘Hume considers experience to be the ultimate source of authority, and it is this reverence for experience that inspires Hume’s foundational ambition’. Hume criticises systems that are not founded on adequate experience as common sources of ‘illusion and mistake’ (EPM 1.10)*.* He also derides the work of ‘the most eminent philosophers’ for lacking evidence (which presumably includes empirical evidence):

Principles taken upon trust, consequences lamely deduced from them, want of coherence in the parts, and *of evidence in the whole*, these are every where to be met with in the systems of the most eminent philosophers, and seem to have drawn disgrace upon philosophy itself. (THN Intro 1, emphasis added)

Similarly, he criticises ‘the moral philosophy transmitted to us by Antiquity’ for ‘being entirely Hypothetical, & depending more upon Invention than Experience’ (HL i.16). More pointedly: ‘none but a fool or madman will ever pretend to dispute the authority of experience’ (EHU 4.20). As Hume makes clear: ‘None of [the sciences and arts] can go beyond experience, or establish any principles which are not founded on that authority’ (THN Intro 10). In contrast to such careless philosophers, Hume describes himself as promising to ‘draw no conclusions but where he is authorized by experience’ (A2).

Explanatory power – that is, how much a theory explains – is also important to Hume. He speaks of rendering our principles *more general and comprehensive* (and therefore more explanatory powerful), pointing out that philosophy is simply ‘a more regular and methodical operation of the same kind’ (DNR 134); he approves of the abstruse philosophers, who ‘proceeding from particular instances to general principles... still push on their enquiries to principles more general, and rest not satisfied till they arrive at those original principles, by which, in every science, all human curiosity must be bounded’ (EHU 1.2); he also endeavours ‘to render all our principles as universal as possible (THN Intro 8). Hume therefore thinks that general principles should predict phenomena not merely at a local but also general level; theories that are limited in their predictions are, *ceteris paribus*, poorer theories. Hume seeks principles that explain and predict as much as possible within a certain set of theoretical constraints; that is, he endorses explanatory power as a theoretical virtue.

Of the three theoretical virtues, Hume says the most about simplicity, syntactical simplicity in particular.[[29]](#footnote-29) He argues that our general principles should be few in number:

…we must endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes… (THN Intro 8)

Hume further emphasises the importance of syntactical simplicity when he claims that we should attempt to explain new phenomena without needlessly multiplying principles; such a wanton postulation indicates the falsehood of a theory:

The antients, tho’ sensible of that maxim, that nature does nothing in vain, contriv’d such intricate systems of the heavens, as seem’d inconsistent with true philosophy, and gave place at last to something more simple and natural. To invent without scruple a new principle to every new phaenomenon, instead of adapting it to the old; to overload our hypotheses with a variety of this kind; are certain proofs, that none of these principles is the just one, and that we only desire, by a number of falsehoods, to cover our ignorance of the truth. (THN 2.1.3.7)

Such a subtility is a clear proof of the falshood, as the contrary simplicity of the truth, of any system. (THN 1.3.16.3)

Not only should our principles be few in number, but they should also be simple in themselves: [[30]](#footnote-30)

...principles... are commonly but few and simple, and that ’tis the sign of an unskilful naturalist to have recourse to a different quality, in order to explain every different operation. (THN 2.1.3.6).

…the utmost effort of human reason is, to reduce the principles, productive of natural phaenomena, to a greater simplicity, and to resolve the many particular effects into a few general causes, by means of reasonings from analogy, experience, and observation. (EHU 4.12)

Hume also describes the ‘simplicity’ of his system as its ‘principal force and beauty’ (THN 2.2.6.2). In short, Hume clearly takes simplicity to be crucial to the acceptability of a theory.[[31]](#footnote-31)

Hazony (2014) persuasively argues that Hume follows Boyle in his emphasis on the reduction and resolution of terms and principles. Notably, Hazony (2014, p.150) argues that Boyle thinks that the merit of a scientific theory is based upon two qualities: simplicity, and generality. As Hazony (2014, p.152) puts it: ‘what makes a scientific theory worth attending to is that it is simpler than its competitors, and that the range of what can be explained by means of it is far greater’. In short, Hume follows Boyle in taking simplicity and what I call explanatory power to be theoretical virtues.

Let us now examine Hume’s conativist thesis through the lenses of these theoretical virtues.[[32]](#footnote-32) He finds his conativist thesis to be simple, explanatorily powerful, and empirically substantiated, and therefore endorses it as a general principle. However, not all of Hume’s data align neatly with his theory; there exist observational outliers, such as the case of the reluctantly assiduous student. Nevertheless, given that Hume’s systematised theory is sufficiently simple and explanatorily powerful, he is justified in providing a conativist explanation of the phenomena that is perhaps less phenomenologically plausible, despite the cost to empirical confirmation.

To highlight the theoretical strength possessed by Hume’s conativist thesis, it is useful to compare Hume’s conativist thesis *C* with the alternative rationalist hypothesis *R*:

***Key****:* Mx – x plays the end-setting role in motivation; Px – x is a passion; Rx – x is reason.

**Hume’s Conativist Hypothesis *C***: ∀x (Mx → Px)

“The end-setting role in motivation is played by the passions.”

**Rationalist Hypothesis *R***: ∀x (Mx → (Px v Rx))

“The end-setting role in motivation is played either by the passions or by reason.”

As is plain from the logical form of the two theses, *C* is syntactically simpler than *R*, since *R* postulates a second end-setter in motivation. Moreover, *C* is also ontologically simpler than *R*. *R* has to admit that there exist end-setting passions, and so *C* does not postulate more than *R* does. However, *R* has to postulate more than *C* in admitting a *sui generis* faculty of practical reason, the existence of which *C* denies. In short, *C* is clearly both syntactically and ontologically simpler than *R*.

With respect to explanatory power, Hume would claim that *C* enjoys the advantage in virtue of the way it flows out of the overall psychological framework of the *Treatise*: Hume explains the calmness of passions partly by appealing to the actions of custom (THN 2.3.4.1), which plays a large role in Book 1, for instance; Hume also uses his general theory of the passions to explain a wide range of behavioural phenomena. However, the rationalist might make a similar claim that *R* fits naturally into rationalist psychology and also explains the relevant phenomena. However, there is one clear respect in which *C* trumps *R* in explanatory power. While *C* postulates only respectable entities, *R* is forced to postulate a somewhat mysterious, unexplained faculty of practical reason. This incurs a cost in explanatory power on *R*, since it is forced to admit an entity which is itself poorly explained – there is something significant that *R* leaves unexplained. The problem is not just that *R* postulates more than *C* (hence why it suffers with regard to simplicity), but that the things it postulates are themselves ill-explained. While *C* and *R* both arguably explain all the relevant human phenomena, *C* does so without appealing to unexplained components, and so *R*’s framework leaves more unexplained that *C*’s; all in all, *C* has the more explanatorily complete theory.

Of course, as granted for the sake of argument, *C* is admittedly less empirically substantiated than *R* precisely because of the apparent counterexamples such as the case of the assiduous student: in such cases, the phenomenology generally suggests that reason plays the end-setting role in motivation (those with a delicacy of taste and passion aside), and therefore the empirical evidence supports *R* more than it does *C*.

However, despite its deficiency in empirical confirmation, one can see why Hume would favour *C* given its advantage in simplicity and explanatory power;[[33]](#footnote-33) Hume is plausibly justified in endorsing *C* even in the face of these apparent counterexamples. Even if one disagrees with Hume’s weighting of the theoretical virtues and thinks empirical confirmation to be vastly more important than simplicity and explanatory power, one should still concede that Hume position seems reasonable enough; he is certainly not being obviously irrational or *ad hoc* in taking this methodological stance. *C* is therefore defensible according to the framework of the theoretical virtues.[[34]](#footnote-34)

### Rules by which to Judge of Causes and Effects

We can also defend Hume’s conativist thesis from the accusation of *ad hockery* by looking to Hume’s ‘rules by which to judge of causes and effects’. Hume offers these rules (which are largely motivated by his analysis of causation as constant conjunction) as guidelines to our causal theorising: ‘Since therefore ’tis possible for all objects to become causes or effects to each other, it may be proper to fix some general rules, by which we may know when they really are so’ (THN 1.3.15.2). Hume’s fourth rule is especially relevant:

The same cause always produces the same effect, and the same effect never arises but from the same cause. This principle we derive from experience, and is the source of most of our philosophical reasonings. (THN 1.3.15.6)

Here Hume puts forth the principle that the same cause always produces the same effect. This principle seems to be patently false at first sight: a fire can be caused by a match or by lightning, for instance. This is uncharitable, and we can save Hume’s blushes by reading this rule as concerning *deepest* causes. In both cases, the fire is ultimately caused by a combination of heat and combustible material. Indeed, Hume sees his project as one concerned with deeper, more general causes, characterising his project as ‘explaining all effects from the simplest and fewest causes’ (THN Intro 8). Hume might claim that finding a plurality of causal explanations for the same effect indicates that we have not yet arrived at ‘original principles’ (EHU 1.2), and must dig deeper; the same effect is always produced by the same *deepest* cause.

In light of this, compare again *C* with *R*. *R* threatens to violate the above principle, as it postulates different causes for multiple instances of the same effect: sometimes the end-setting role in motivation is played by reason, and other times it is played by the passions.[[35]](#footnote-35) Importantly, it is difficult to see reason and passions sharing a deeper mutual cause that would explain the end-setting aspect of motivation, which proponents of *R* would have to provide if they are not to fall foul of Hume’ fourth rule; at the very least, *R* suffers a theoretical cost by light of Hume’s rules without some such deeper explanation. And indeed, the rationalist proponents of *R* would be loathe to offer such an explanation on pain of removing reason from the pedestal that it occupies in their framework.

*C*, however, satisfies this principle, as it postulates a single cause for instances of the same effect; in comparison to *R*, *C* clearly has the theoretical advantage by light of Hume’s rules.[[36]](#footnote-36) Notably, Hume takes the fourth rule to be ‘the source of most of our philosophical reasoning’; thus, it is extremely plausible that he was explicitly guided by this principle in evaluating *C*. In short, Hume’s theory of motivation, far from being *ad hoc*, is in fact a shining example of his philosophical methodology in action.

### Reflective Equilibrium and Reinterpreting Data

Once a theory is accepted, we reinterpret data in light of this theory; sometimes reflective equilibrium is attained by favouring general principles over particular instances. Recall Demeter’s (2012) characterisation of Hume’s experimental method as one consisting in a kind of analysis and synthesis. We use phenomena to derive general principles (analysis), and then use these principles to explain phenomena (synthesis). But not all phenomena will be explained by the theory, and if our theory is strong enough – that is, if it explanatorily powerful and simple, in line with Hazony’s (2014) reading of Boyle and Hume – we are licensed to explain *away* troublesome phenomena within the framework of this theory in our synthesis. Once Hume rules in favour of his theory of motivation, he can now use his systematised theory to examine the recalcitrant observations through new theoretical lenses. He concludes that in these cases, there is an unnoticed calm passion at work, despite our pre-theoretic intuitions to the contrary. This judgment is justified given the theoretical soundness possessed by Hume’s theory of motivation; accusations of *ad hockery* are, if not refuted, at the very least mitigated by these theoretical considerations. Therefore, Hume can justifiably postulate the calm passions in explaining away purported counterexamples to his theory of motivation.

Indeed, Hume does something very similar with the ‘monkish virtues’ in the second *Enquiry* (EPM 9.3) – having derived his definition of virtue as qualities ‘useful or agreeable to ourselves or others’ from his observations of the personal qualities that we praise and blame, he dismisses the ‘monkish virtues’ as vices despite the fact that some might pre-philosophically think at least some of them to be virtues (humility, for example, is still widely considered a virtue today), because he finds them to be disagreeable and detrimental to the self and to others:

And as every quality, which is useful or agreeable to ourselves or others, is, in common life, allowed to be a part of personal merit; so no other will ever be received, where men judge of things by their natural, unprejudiced reason, without the delusive glosses of superstition and false religion... the whole train of monkish virtues; for what reason are they every where rejected by men of sense, but because they... cross all these desirable ends; stupify the understanding and harden the heart, obscure the fancy and sour the temper. We justly, therefore, transfer them to the opposite column, and place them in the catalogue of vices; nor has any superstition force sufficient among men of the world, to pervert entirely these natural sentiments. (EPM 9.3)

Hume says that ‘we justly, therefore, *transfer’* (emphasis added) the monkish virtues to ‘the catalogue of vices’: this indicates that he thinks that they are pre-theoretically treated as virtues (hence why a ‘transfer’ is necessary),[[37]](#footnote-37) and that his general principles give sufficient reason to abandon this intuition.[[38]](#footnote-38) Here Hume is favouring his general theory over the particular instances, because he finds it to be theoretically well-founded; in this case, reflective equilibrium is attained by dismissing the latter to accommodate the former. This is directly analogous to what Hume does with respect to the calm passions, and in neither case is Hume theorising in an obviously *ad hoc* manner.

## Conclusion

In this paper I hope to have shown how an empirical Hume’s account of the calm passions may be theoretically defended, and why conceptual accounts would prove problematic. This is not a full treatment of the complex issues of the calm passions, and does not pretend to be. Notably, this paper does not discuss non-conceptual defences of Hume’s appeal to the calm passions, nor does it address in detail Hume’s discussion of the calm/violent distinction.

Nevertheless, this paper hopes to have both historical and contemporary significance. From a historical perspective, I have hoped to shed some light on Hume’s theoretical framework, both explicit (his rules by which to judge of causes and effects) and implicit (the theoretical virtues). In using these frameworks to rule out rationalist theories of motivation, it becomes apparent just how important simplicity is to Hume, and how restrictive his fourth rule by which to judge of causes and effects really is.[[39]](#footnote-39) Further work could be done on how exactly these theoretical frameworks (particularly his emphasis on simplicity and the fourth rule) may be wielded to dismiss other hypotheses Hume objects to. For instance, it might prove interesting to investigate to what extent these frameworks might prove useful in illuminating his dismissal of the Design Argument for God’s existence (Section 11 of the *Enquiry* and the *Dialogues*).

From a contemporary perspective, I have hoped to have gone a small way in showing that by appealing to the simplicity of conativism, as well as using its framework to provide plausible explanations for apparently problematic phenomena, conativism can reasonably be defended not just as a conceptual teleological thesis (e.g. Smith 1994), but as an empirical causal thesis. Sinhababu (forthcoming) carefully does much of this careful legwork, defending a modern conativist account (construed as a psychological thesis) from attack, in part by appealing to theoretical considerations not dissimilar to the ones discussed in this paper, and in part by using its framework to explain an extremely wide range of human phenomena such as *akrasia*, willpower, and Gendler’s ‘aliefs’ (2008). Sinhababu sees his account to be strongly inspired by Hume; I hope to have shown that he is right to read Hume in this light, and that Hume in turn was right in offering his original account in the first place.[[40]](#footnote-40)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## ­­­­­­­­­­References

Baillie, James, 2000. *Hume on Morality*, Routledge.

Boehm, Miren, 2013. “Hume’s Foundational Project in the *Treatise*’, *European Journal of Philosophy* (Early View). doi: 10.1111/ejop.12056.

Top of Form

Bricke, John, 1996. *Mind and Morality: An Examination of Hume’s Moral Psychology*, OUP.

Cohon, Rachel, 2008. *Hume’s Morality: Feeling and Fabrication*, OUP.

Demeter, Tamás, 2012. “Hume’s Experimental Method”, *British Journal for the History of Philosophy* 20:3.

Eells, Ellery, 1991. *Probabilistic Causality*, Cambridge, U.K.: Cambridge University Press.

Gendler, Tamar Szabó, 2008. Alief and Belief. *Journal of Philosophy* 105 (10): 634-663.

Goodman, Nelson, 1961. “Safety, Strength, Simplicity”, Philosophy of Science 28 (2):150-151.

Goodman, Nelson, 1965. *Fact, Fiction, and Forecast*, Bobbs-Merrill.

Hazony, Yoram, 2014. “Newton’s Explanatory Reduction and Hume’s System of the Sciences”, in Biener, Z., and Schliesser, E. (eds.), *Newton and Empiricism*, Oxford University Press.

Hazony, Yoram, and Schliesser, Eric, Forthcoming. “Newton and Hume”, in Russell, P. (ed.), *The Oxford Handbook of Hume*, Oxford University Press.

Hume, David, 1739-1740. A Treatise of Human Nature, edited by L. A. Selby-Bigge, 2nd ed. revised by P.H. Nidditch, Oxford 1975.

Hume, David, 1748. Enquiries concerning Human Understanding and concerning the Principles of Morals, edited by L. A. Selby-Bigge, 3rd edition revised by P. H. Nidditch, Oxford 1975.

Hume, David, 1777. *Essays Moral, Political and Literary*, rev. edition edited by E. Miller, Liberty Fund Inc., 1985.

Hume, David, 1779. *Dialogues Concerning Natural Religion*, edited by N. Kemp Smith, Bobbs-Merill Educational Publishing, 1947.

Hume, David. *The Letters of David Hume*, edited by J.Y.T Greig, OUP 1932.

Immerwahr, John, 1991. “The Anatomist and the Painter: The Continuity of Hume’s *Treatise* and *Essays*, *Hume Studies* 17.1: 1-14.

Kant, Immanuel, 1781/1787. *Critique of Pure Reason*, translated and edited by Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1997).

Lewis, David, 1973a. Counterfactuals, Oxford: Basil Blackwell.

Lewis, David, 1973b. “Causation”, *Journal of Philosophy* 70: 556–67.

Lewis, David, 1986. *On the Plurality of Worlds*, Oxford: Blackwell Publishers.

Loeb, Louis E., 2002. *Stability and Justification in Hume's* Treatise, Oxford University Press.

Mackie, J.L., 1965. “Causes and Conditions”, *American Philosophical Quarterly*, 12: 245–65.

Marusic, Jennifer S., 2010. ‘Does Hume Hold a Dispositional Account of Belief?’. *Canadian Journal of Philosophy* 40:2.

McMullin, E., 2008. “The Virtues of a Good Theory”, in Psillos, S., and Curd, M. (eds.), *The Routledge Companion to Philosophy of Science*, Routledge.

Phillos, S., 2005. *Scientific Realism: How Science Tracks Truth*, Routledge.

Qu, H., 2012. “The Simple Duality: Humean Passions”, *Canadian Journal of Philosophy*, 42:supp1, 92-116.

Qu, H., 2016. “Prescription, Description, and Hume’s Experimental Method”, *British Journal of the History of Philosophy*.

Qu, H., Forthcoming. “Hume on Mental Transparency”, Pacific Philosophical Quarterly.

Schliesser, E., 2007. “Hume’s Newtonianism and Anti-Newtonianism”, *Stanford Encyclopedia of Philosophy*.

Schliesser, E., 2010. “Copernican Revolutions Revisited in Adam Smith by way of David Hume”, *Revista Empresa y Humanismo* 13:213-248.

Shaw, Daniel, 1989. “Hume’s Theory of Motivation”, *Hume Studies* 15:1.

Sinhababu, Neil, forthcoming. *A Treatise of Humean Nature*, Oxford University Press.

Smith, Michael, 1994. *The Moral Problem*, Blackwell.

Stroud, Barry, 1977. *Hume*, Routledge.

Zellner, A., Keuzenkamp, H. & McAleer, M. (eds.), 2001. *Simplicity, Inference and Modelling: Keeping It Sophisticatedly Simple*, Cambridge: Cambridge University Press.

1. The characterisation of Hume’s account of reason is a notoriously contentious issue which I cannot adequately address due to space constraints. Here I only make the uncontroversial claim that reason is the faculty that concerns the generation of beliefs in some way: ‘Reason is the discovery of truth and falsehood’ (THN 3.1.1.9), and since Hume thinks that only beliefs are truth-apt (c.f. THN 2.3.3.5), it seems clear that reason at the very least concerns the production of beliefs. [↑](#footnote-ref-1)
2. Although reason can work in conjunction with the passions to generate new passions, which in turn can play the end-setting role in motivation. I have a desire for sweet things, reason tells me that the ice-cream is sweet, and this produces a passion for the ice-cream, which motivates me to reach for it. [↑](#footnote-ref-2)
3. In the references to Hume’s texts throughout, ‘THN’ refers to the *Treatise of Human Nature*, ‘A’ to the *Abstract of a Book Lately Published*, ‘EHU’ to the *Enquiry Concerning Human Understanding*, ‘EPM’ to the *Enquiry Concerning the Principles of Morals*, ‘EMPL’ to *Essays Moral, Political, and Literary,* ‘DNR’ to the *Dialogues Concerning Natural Religion*’, and ‘HL’ to *The Letters of David Hume*. Arabic numerals refer to section and paragraph numbers (EHU and EPM); book, part, section, and paragraph numbers (THN); or to paragraph numbers (A). EMPL numbers refer to pages in the Miller revised edition of the *Essays* (Liberty Fund Inc., 1985), and DNR numbers to pages in the Kemp Smith edition of the *Dialogues* (Bobbs-Merill Educational Publishing, 1947). HL Roman numerals refer to volume, and Arabic numerals to page numbers in the Greig edition of the *Letters* (OUP, 1932). [↑](#footnote-ref-3)
4. The calm passions are not ever-present but rather occurrent; indeed Hume argues that there is no impression that is constant and invariant throughout our lives (THN 1.4.6.2). Detailing their triggers is a difficult task which is in any case tangential to my purposes, thus I will abstain from doing so here. [↑](#footnote-ref-4)
5. Hume certainly thinks that calm passions can possess more causal efficacy than their violent counterparts; what determines their relative causal efficacies is the agent’s ‘strength of mind’ (THN 2.3.3.10). [↑](#footnote-ref-5)
6. A different worry is that such an introspective mistake would violate Hume’s Incorrigibility Principle (c.f. THN 1.4.2.7). I lack the space to adequately address this worry here, but in my Qu (Forthcoming) I argue that Hume’s Incorrigibility Principle only bears on first-order qualitative character and not on higher-order mental states; a mistaken second-order belief about our passions (such as mistaking them for the actions of reason) is allowed by this principle. And indeed, the principle would be an implausible one if it did not allow for errors of this kind. [↑](#footnote-ref-6)
7. Briefly, functionalist theories of the mind characterise mental states by their functional role, rather than, say, their qualitative character. [↑](#footnote-ref-7)
8. Here I assume that desires are, by definition, passions. Stroud therefore subscribes to a more restricted form of conativism than Hume (that is, those passions that are desires perform the end-setting role in motivation). [↑](#footnote-ref-8)
9. Note that Smith (1994, p.114), like Stroud, also sees his account of the desires as dispositional. Note also that Smith’s and Bricke’s accounts, like Stroud’s, deals with desires rather than passions more generally; correspondingly, Bricke’s interpretation gives desires a pride of place in that all passions ultimately presuppose desire and/or aversion (p.36). [↑](#footnote-ref-9)
10. Smith (1994, p.103) emphasises that his thesis is not a causal one, but a teleological one in giving an explanation for action; Bricke (1996, p.8) meanwhile treats the conativist thesis causally. My paper is, like Hume, concerned with the causal version, and therefore does not directly address Smith’s framework, but rather Bricke’s appropriation of it. Moreover, Smith does not take his account to be an interpretation, but a thesis that is merely Humean in spirit, whereas Bricke attributes the thesis to Hume. Admittedly, to what extent Bricke’s account is a straightforward interpretation is uncertain: he talks of ‘regimenting’ Hume’s texts and revising his theories, and explicitly abandons Hume’s psychological atomism despite attributing the thesis to him. Bricke (p.25) also describes his account as only ‘insecurely Hume’s’. Bricke’s account therefore seems more in the spirit of an interpretive rational reconstruction, teasing out a comprehensive, cohesive account from Hume’s less sophisticated theory. In rejecting this account I argue that even a reconstruction of this nature fares worse than Hume’s original account, and that Hume would have resisted this reconstruction for reasons above and beyond the denial of his psychological atomism. [↑](#footnote-ref-10)
11. I cannot think of any modern theory of causation that treats the relation as metaphysically necessary; Lewis’(1973b) counterfactual analysis, Mackie’s (1965) INUS conditions analysis, and probabilistic accounts (e.g. Eells 1991) all take causation to be contingent. [↑](#footnote-ref-11)
12. These are of course not the only interpretive problems with a conceptual reading: such a reading would contradict Hume’s impressionistic theory of the passions, as well as their indefinability (THN 2.1.2.1); furthermore, it would be difficult to make sense of such a reading from within Hume’s overall project of providing an *empirical* science of man. Here I focus on the interpretive problems which are also substantive philosophical worries. [↑](#footnote-ref-12)
13. Arguably, the claim that ‘the Higgs-Boson, if it exists, is capable of granting mass’ might be necessarily true, but this is not the interesting claim, which is that the Higgs-Boson *actually* grants mass. Similarly, conativism is interesting only if treated as a thesis about what actually does and does not motivate in an end-setting role, rather than what, if existent, is capable of doing so. [↑](#footnote-ref-13)
14. Bricke arguably sees his account as an empirical one, but his treatment of evaluative beliefs casts doubt on whether this view is justified: if his account were really an empirical one, he would not be licensed to rule that evaluative beliefs are really desires despite their empirically better matching the description of beliefs *simply because* they play the end-setting role in motivation (Bricke 1996, p.28). Similarly, Stroud’s account cannot be empirical, because he characterises desires as bare dispositions to perform the end-setting role in motivation (Stroud 1977, pp.167-8). [↑](#footnote-ref-14)
15. I remain neutral on whether Hume’s account provides an empirical functional characterisation (as opposed to a phenomenal characterisation) of the passions, independently of worries of *ad hockery*. [↑](#footnote-ref-15)
16. Loeb (2002) argues that Hume is fundamentally committed to a dispositional account of beliefs, although his official theory treats beliefs as occurrent; vivacity would correspondingly be treated as something like firmness or steadiness rather than as something phenomenological. I find this account implausible as an interpretation of Hume. For a convincing response to Loeb’s interpretation of Hume on beliefs, see Marusic (2010). [↑](#footnote-ref-16)
17. The discussions of systematisation, the theoretical virtues, and reflective equilibrium is something I also make use of in Qu (2016), in exploring how Hume derives genuine prescriptions from his seemingly descriptive experimental method. [↑](#footnote-ref-17)
18. Indeed, lack of systematisation is precisely what separates bad general rules from good: the latter reveal the ‘irregular nature’ of the former, which are ‘destructive of all the most establish’d principles of reasoning’ (THN 1.3.13.12). [↑](#footnote-ref-18)
19. Kant’s notion of systematicity, although better-developed, is fundamentally similar, involving the subsuming of knowledge under fewer and fewer principles, as he notes in the *Critique of Pure Reason* (B674). The key difference is that for Kant, the rules for systematisation are *a priori*, whereas Hume’s criteria are *a posteriori*; Kant’s notion of systematicity is also deeply hierarchical in a way that Hume’s is not. [↑](#footnote-ref-19)
20. Hume uses the term ‘principle’ in two senses: as a type of cause or origin, and as a general rule. There seems to be some equivocation between the two senses in this passage: clearly Hume’s reference to finding principles that ‘excite our sentiments’ makes use of the former sense (since general rules do not often do so), but his speaking of proceeding from particular instances to general principles clearly makes use of of the latter sense. [↑](#footnote-ref-20)
21. This should not be read as denying that the passions have intentionality. For discussion of how to reconcile the intentionality of the passions with their qualitative character and also their simplicity, see Qu (2012). [↑](#footnote-ref-21)
22. I see Stroud’s (1977, pp.158-61) reading of this passage as conceptual, turning as it does on the nature of propositions. Bricke’s (1996) reading of the argument also seems conceptual, as the following discussion of the original existences argument makes clear: ‘Reflection on the conditions on acting for a reason, and of explaining an action by citing the agent’s reason for acting, when deepened by reflection on the concept of direction of fit, leads directly to the conativist conception of reasons for action. (Perhaps that is why Hume neglected to spell out the steps.)’ (Bricke 1996, p.28). Here, *a priori* reflection on our concepts suffices for the conclusion, which certainly indicates a conceptual argument. [↑](#footnote-ref-22)
23. These theoretical virtues are not meant to be an exhaustive list, although they are the three that I think Hume to recognise, and also the minimum needed for an adequate set of theoretical criteria. We need empirical confirmation, since a theory cannot so much as get off the ground if it fails to explain our empirical observations. But this is insufficient, as many different theories will account for the empirical phenomena; we also need other criteria, such as simplicity. We cannot have simplicity without explanatory power or vice versa, since they counterbalance each other. If we cared only about simplicity, we would only have very austere theories that did not make many predictions. If we cared only about explanatory power, we would have overly extravagant theories that wantonly multiplied principles (the sign of an ‘unskilful naturalist’ [THN 2.1.3.6]) in their quest to predict everything. Of course simplicity and explanatory power need not always conflict; sometimes a simpler theory might turn out to be more explanatorily powerful as well. In any case, as a basic conception these three virtues perform adequately in theory selection. This is similar to Goodman (1961), who advocates simplicity, strength (that is, explanatory power), and safety (which we can take as corresponding to empirical confirmation) as theoretical virtues. Such a taxonomy is of course cruder than contemporary accounts of theoretical virtues (see McMullin [2008] for an excellent discussion of various types of theoretical virtues), but this would not be surprising from an 18th century thinker such as Hume. [↑](#footnote-ref-23)
24. I am using ‘evidence’ in a non-factive manner. I also include pre-philosophical intuitions under ‘observed evidence’, given that some of Hume’s data will source from these intuitions; for example, in the second *Enquiry*, Hume examines our pre-philosophical intuitions regarding virtue and vice in deriving his account of morality. [↑](#footnote-ref-24)
25. I use ‘explanatory power’ in a narrower fashion than commentators such as Psillos (2005, p.165), who take it to encompass a variety of theoretical virtues, including unifying power (which ties in with simplicity). McMullin (2008, p.502) takes a similar view of explanatory power to Psillos. [↑](#footnote-ref-25)
26. Simplicity has received much attention in the philosophy of science. For a multi-disciplinary discussion, see the papers in Zellner, A., Keuzenkamp, H. & McAleer, M. (eds.), 2009. [↑](#footnote-ref-26)
27. These two forms of simplicity often come into conflict: sometimes it might be that positing extra entities allows us to simplify our general principles. For example, postulating Neptune allowed astronomers to explain perturbations in the orbits of other planets without needlessly complicating astronomical principles. Lewis (1986) famously was happy to sacrifice quantitative parsimony for qualitative in postulating an infinitude of possible worlds in order to eliminate abstract entities such as properties and propositions. [↑](#footnote-ref-27)
28. It should be noted that syntactic simplicity is an empty notion if language is up for grabs, as Goodman (1965) notes. Syntactic simplicity is always relative to a language, and so we need to fix our language before examining relative syntactic simplicity. [↑](#footnote-ref-28)
29. Schliesser (2010) contains some interesting discussion of Hume on simplicity and beauty with respect to Copernicus. [↑](#footnote-ref-29)
30. Indeed, there is no point minimising the number of principles without paying any attention to the simplicity of these principles. A theory consisting of a single principle which was a conjunction of an arbitrarily large number of propositions would certainly not count as simple, for instance. [↑](#footnote-ref-30)
31. It is important to note that despite his emphasis on simplicity, Hume does not take it to be an *overriding* consideration. He recognises that theories can easily run into falsehoods when they single-mindedly pursue simplicity at the expense of the other theoretical virtues: in arguing against the thesis that men are solely motivated by self-love, he derides ‘that love of simplicity, which has been the source of much false reasoning in philosophy’ (EPM App 2.6). This is not to say that he denies the importance of simplicity, only that we should be careful in pursuing it: ‘And though they [moralists] have sometimes carried the matter too far, by their passion for some one general principle; it must, however, be confessed, that they are excusable in expecting to find some general principles, into which all the vices and virtues were justly to be resolved’ (EHU 1.15). Simplicity has to be weighed up against empirical confirmation and explanatory power in determining the worth of a theory. [↑](#footnote-ref-31)
32. We would expect a theory of motivation to be simple and explanatorily powerful only if we presumed that motivation is a *natural kind*, that is, if motivation is projectible, grounds counterfactuals, and so forth. If motivation were a gerrymandered concept, an adequate analysis of it might turn out to be very complex and explanatorily weak. It seems clear from Hume’s ambitions of introducing the experimental method to moral subjects (as stated in the subtitle of the *Treatise*) and providing a science of man (THN Intro 4) that he thinks that motivation (among the other mental phenomena he discusses) is a natural kind in the relevant sense. [↑](#footnote-ref-32)
33. Note that I am not claiming that *C* should be favoured simply because it wins ‘two to one’; the weighting process will likely have to be far more complex than this. There are no hard and fast rules as to how conflicts between theoretical virtues should be decided. [↑](#footnote-ref-33)
34. The above reasoning holds for the comparison of *C* with *R*, but what about the simpler hyper-rationalist hypothesis *H* that postulates reason as the end-setter in *every* case of motivation? *H* Hume has ruled out already: he argues in THN 2.3.3.2-3 (prior to his discussion of the calm passions) that there are many clear-cut cases whereby reason is *not* the end-setter in motivation; the problem cases for *C* on the other hand are less numerous, and it is not entirely clear that the end-setter in these cases is unambiguously reason. Although *H* is approximately as simple as *C*, it fails with respects to empirical confirmation, and is therefore Pareto dominated by *C* with respect to theoretical soundness. [↑](#footnote-ref-34)
35. Perhaps one could defend *R* by maintaining that we should individuate effects finely enough such that ‘being motivated by a passion’ counts as a different effect from ‘being motivated by reason’, and so the principle ‘same effect, same cause’ does not apply here. However, there does not seem to be a principled reason for distinguishing these two effects, as the behaviour of the agents in both cases does not seem as though it should relevantly differ. [↑](#footnote-ref-35)
36. Thesis *H* also satisfies Hume’s principle, but as noted in an earlier footnote, Hume rules it out early on in his argument. Presumably he would think it to violate his third rule (THN 1.3.15.7) requiring constant conjunction between cause and effect, given that he argues in THN 2.3.3.2-3 that reasoning is frequently observed to fail to correspond to an end-setting state, and only does so when there is a passion present (which is itself the true end-setter). Of course, the apparent counterexamples such as the assiduous student also threaten *C* with respect to the third rule, but Hume presumably thinks the passions to be *more obviously* constantly conjoined in an end-setting role with motivation than reason seems to be (which is not very much at all, if we take his argument in THN 2.3.3.2-3 seriously); as mentioned earlier, the problem cases for *C* are more ambiguous than the problem cases for *H*, and *C* seems to win on this count. In this respect, Hume’s treatment of *H* is analogous to his discussion of the claim that we are only motivated by self-love (EPM App 2); this hypothesis satisfies Hume’s fourth rule, but he dismisses it because of the various counterexamples he finds (which entail that it violates the third rule). [↑](#footnote-ref-36)
37. Hume does note that the monkish virtues are rejected by ‘men of sense’, but of course such men may be rare; certainly in Hume’s time the monkish virtues were commonly taken as virtues, and even today humility is widely considered a virtue. The ‘men of sense’ who reject the monkish virtues can be seen as analogous to the discerning agents who successfully discern a calm passion in cases of motivation such as the assiduous student. [↑](#footnote-ref-37)
38. Admittedly, this might not be a purely theoretical correction, since the transfer is at least partially the result of our ‘natural sentiments’, which are not ‘entirely’ perverted. Nevertheless, it seems clear that theoretical considerations can and do play a role in pushing us one way or another, especially when our sentiments are on the fence (because only partially perverted). [↑](#footnote-ref-38)
39. Similar remarks might be made regarding some of Hume’s other rules by which to judge of causes and effects. For instance, Hazony and Schliesser (forthcoming) argue that Hume’s seventh rule obliquely critiques Newton’s theory of gravitation. [↑](#footnote-ref-39)
40. For invaluable and detailed comments on this paper, I am indebted to Don Garrett. I owe much also to Béatrice Longuenesse, Peter Millican, and Kelley Schiffman for very helpful feedback. Thanks also to Harjit Bhogal, Camil Golub, and Alexis Elder for useful comments on an earlier draft. I would also like to thank an audience at the Yale-UConn Graduate Philosophy Conference 2012 for their excellent comments. [↑](#footnote-ref-40)