This unpaginated preview of Yuhei Yoshioka's "Hume's Calm Passions and Motivation: A Defence of the Occurrent State Interpretation" is based on a preliminary draft that has not been copyedited or typeset and that does not reflect the author's final corrections. It is temporarily provided for preliminary information only and should not be cited, copied, retained, or redistributed without the author's permission.

The citeable version of record will appear on Project Muse in Hume Studies 52(2) in spring of 2027.

© 2027 Yuhei Yoshioka
c121670k-yuhei@hiroshima-u.ac.jp

Abstract: Hume, who argued that reason alone cannot produce action, explained cases in which we appear to act in accordance with reason by appealing to "the calm passions." Because Hume provided no explicit grounds for this explanation, scholars remain divided over their nature and motivational role. This paper argues that the calm passions relevant to motivation are best understood as occurrent rather than dispositional states. More specifically, it claims that, with the exception of "the general appetite to good, and aversion to evil, consider'd merely as such," these passions should be understood as occurrent passions produced by instinct. To support this claim, the paper examines how these passions arise, focusing on the only passion of this kind that Hume discusses in detail, benevolence, and analyses Hume's use of the term "instinct."

Keywords: Hume, David; passions; calm passions; motivation; theory of action; instinct; benevolence; disposition; occurrent state

Introduction

Hume's position on the calm passions plays a crucial role in his theory of motivation. As is well known, he argued that reason alone cannot produce action (T 2.3.3.4).[1] Yet ordinary experience suggests that there are many cases in which we seem to act in accordance with reason. There is thus a marked tension between Hume's theoretical claim and our everyday experience, and the calm passions are key to resolving it. For, according to Hume, even in cases where we seem to act in accordance with reason, this is no more than an appearance: what is at work, and what serves as the motive,[2] is a calm passion that is not readily distinguishable by immediate feeling (T 2.3.3.8). As examples of such calm passions, Hume cited benevolence, resentment, the love of life, kindness to children, and "the general appetite to good, and aversion to evil, consider'd merely as such" (T 2.3.3.8; hereafter, "the general appetite").[3]

Hume, however, offered no explicit grounds for this argument, and interpreters remain divided over what sort of mental states the calm passions are and how they motivate us. As I explain in Section 1, the existing literature has advanced two influential interpretations. One treats the calm passions as occurrent mental states (Árdal 1966; Fieser 1992; Immerwahr 1992;

[1] See the Bibliographical Note to this issue of *Hume Studies* for a key to references to Hume's works.
[2] Especially when discussing the artificial virtues in Book 3 of the *Treatise*, Hume used the term "motive" in a fairly broad sense. In Book 2, however, which is the principal focus of this paper, he used it in a clearly causal and psychological sense, as Cohon (2008, 30–1) points out. Accordingly, following her interpretation, this paper defines a "motive" as the psychological and proximal cause of an action.
[3] Hume frequently used "good" and "pleasure," and "evil" and "pain," as interchangeable concepts (for example, T 2.1.1.4, 2.3.1.1, 2.3.9.8). This paper likewise uses "good" and "pleasure," and "evil" and "pain," interchangeably.

Draft. Do not Cite.

Qu 2018), whereas the other treats them as dispositional states (Stroud 1977; Enç 1996; Shaw 1998; Loeb 2002; Radcliffe 2018). I shall refer to the former as the occurrent state interpretation and to the latter as the dispositional interpretation. Although the former aligns well with the overall theoretical framework of Hume's theory of the passions, it faces the difficulty of explaining how Hume's appeal to the calm passions avoids the charge of being ad hoc, since it leaves the empirical basis of their motivational role obscure. The latter, by contrast, better explains cases in which we seem to act in accordance with reason, but it raises the further question of how passions, understood as dispositions, are to be accommodated within Hume's system. To gain a clearer understanding of Hume's theory of motivation, it is necessary to determine how at least some of the calm passions discussed in connection with motivation are to be understood in light of these competing interpretations.

This paper defends the occurrent state interpretation through a critical examination of the dispositional interpretation, in particular, of Elizabeth Radcliffe's reading of the calm passions as innate dispositions. More specifically, by examining how the calm passions relevant to motivation arise, with particular attention to benevolence, this paper argues that, with the exception of the general appetite, these passions should be taken to be occurrent passions produced by instinct.

This paper makes three contributions. First, in the interpretive debate over the calm passions, it refines the occurrent state interpretation by situating it within the wider framework of Hume's philosophical system and presents it as a more textually persuasive reading.[4] Second, it offers a critical examination of Radcliffe's interpretation, which is regarded as one of the most influential recent readings of Hume's theory of action, and advances an interpretation that is more closely grounded in the texts. Third, from the standpoint of the occurrent state interpretation, it shows that Hume's own discussion should be understood not merely as an ontological claim, but also as offering, at least in the case of benevolence, a psychological explanation of its occurrence.

The structure of this paper is as follows. Section 1 outlines the occurrent state interpretation and the dispositional interpretation and identifies the difficulties facing each. Section 2 presents Radcliffe's interpretation as a more refined version of the dispositional interpretation. Section 3 examines existing criticisms of the kind of interpretation represented by Radcliffe and the limits of those criticisms before developing this paper's own critique. Section 4 sets out my interpretation through an analysis of benevolence and the concept of instinct and clarifies its advantages. Finally, Section 5 draws out the implications of my interpretation for Hume's theory of passions and motivation.

Before proceeding, I should clarify the scope of the calm passions discussed in this paper. For example, Hume presents the sense of beauty and deformity as a paradigmatic case of the calm passions (T 2.1.1.3) and suggests that the moral sentiments are also included among them (T 3.1.2.1). The former, however, is not directly relevant here, while in the latter case the relation to motivation remains unclear and is disputed among scholars. Accordingly, this paper confines its discussion to those calm passions that Hume explicitly discusses in connection with motivation.

1. The Occurrent State Interpretation and the Dispositional Interpretation

This section examines the occurrent state interpretation and the dispositional interpretation. To do so, it begins by explaining Hume's distinction between violent and calm passions. According

[4] Qu (2018) offers a recent defense of an occurrent reading of the calm passions. His concern, however, differs from mine in important respects. He defends an empirical reading of the calm passions primarily at the methodological level without analysing particular calm passions in detail, whereas the present paper focuses on the psychology of particular calm passions through close analysis of the relevant texts.

Draft. Do not Cite.

to Hume, violent passions produce sensible emotion (T 2.3.3.9). By contrast, he characterized calm passions in the following terms:

> Now 'tis certain, 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. These desires are of two kinds; either certain instincts originally implanted in our natures, such as benevolence and resentment, the love of life, and kindness to children; or the general appetite to good, and aversion to evil, consider'd merely as such. 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. (T 2.3.3.8) [5]

On the basis of this contrast, one might interpret the difference between the two as lying not in their ontological status, but in the phenomenological intensity of emotion. This underlies the occurrent state interpretation, according to which the calm passions are occurrent mental states that are nonetheless not readily distinguishable by immediate feeling (Árdal 1966, 93–108; Fieser 1992, 6–7; Immerwahr 1992, 294; Qu 2018, 451–4).

This interpretation is plausible at first glance, but its difficulties emerge when it is considered in the context of motivation. As the passage suggests, Hume took the effects of calm passions to be evidence for their presence. In the present context, the relevant effects include actions. But when an action occurs and no calm passion is readily distinguishable by immediate feeling, the occurrence of that action does not by itself justify treating it as evidence for the presence of a calm passion. Yet Hume does not appear to have provided an explicit independent basis for treating it that way in his discussions of the passions and motivation. On the occurrent state interpretation, then, it remains unclear what warrants supposing that an occurrent calm passion is present in such cases. I shall refer to this problem as "the difficulty of justifying the presence of the calm passions" (hereafter, "the justificatory difficulty").

Of course, one possible response is simply to reject Hume's argument as a failed argument. There is, however, another line of interpretation that seeks to avoid the impasse facing the occurrent state interpretation and to reconstruct Hume's view in a more coherent form. The most representative version of this is the dispositional interpretation, which takes the calm passions to be dispositions (Stroud 1977, 154–70; Enç 1996, 240–3; Shaw 1998, 55–79; Loeb 2002, 3–6; Radcliffe 2018, 14–64).

In examining this interpretation, I begin with Barry Stroud's reading. According to him, the calm passions can be understood as dispositions by virtue of which a person is moved to perform a certain action when she believes that the action will bring about a certain result (168).[6] Indeed, taking kindness to children as a concrete example of the calm passions, Stroud analysed it as a disposition to be moved, given the relevant belief, to treat children kindly or to protect them from harm (165). Such a disposition can itself serve as a motive for action, and because it makes no reference to emotion, it motivates without producing any perceptible feeling. This interpretation is therefore consistent with the distinctive feature of the calm passions.

If Stroud's reading is adopted, a new response to the justificatory difficulty becomes available. According to Qu (2018, 446), this reading conceptually defines the calm passions as states whose function in motivation is to set a given action as an end. On this understanding,

---

[5] Although Hume here wrote of "tendencies," the term need not, by itself, be taken to mark a distinct dispositional category. He immediately characterized these same states as "real passions" and said that they "produce little emotion," not that they produce none at all. As Qu (2018, 452) notes, this suggests that calm passions are faintly felt rather than non-occurrent.

[6] Stroud took Hume's argument to be defective and presented his dispositional interpretation as a way of remedying that defect. His primary concern, therefore, is not to provide a strictly textual interpretation, but rather to secure theoretical coherence.

Draft. Do not Cite.

there is no need to establish independently through perceptual experience that a further occurrent passion is present behind the action, since the relevant passion is identified in terms of the dispositional role it plays in motivation. The question of whether the calm passions serve as motives is therefore reduced to the adequacy of this conceptual definition.

Although Stroud's interpretation has certain advantages, it also gives rise to a new problem. According to Qu (447–8), if this interpretation is correct, Hume offers only a trivial account of motivation. On this view, because the calm passions are defined in terms of their end-setting function, they necessarily count as the motives of all actions that lack sensible emotion. But such an account is circular and uninformative and contributes nothing to the explanation of motivation in such actions. Accordingly, even on Stroud's reading, Hume's discussion of motivation by the calm passions remains unsupported. Moreover, because this interpretation understands the calm passions to be dispositional states conceptually defined by their end-setting function, it offers no distinct answer to an important question for this paper, namely, by what psychological mechanism such passions arise.

2. The Instinctive Interpretation

Although Stroud's interpretation leads to difficulties like those facing the occurrent state interpretation, this does not show that the dispositional approach itself is untenable. If one focuses on the fact that Hume characterizes some of the calm passions as our instincts (T 2.3.3.8), this allows for a different version of the dispositional interpretation (Enç 1996, 240; Shaw 1998, 71–5; Loeb 2002, 5; Radcliffe 2018, 14–64; see also Cohon 2008, 38–40). In what follows, I shall call this kind of reading the instinctive interpretation, and examine it primarily through the work of Radcliffe, who offers the most detailed account of it in the existing literature.

The instinctive interpretation understands the calm passions to be innate dispositions that produce occurrent passions in certain circumstances (Radcliffe 2018, 23–4). For example, Radcliffe suggests that benevolence as an innate disposition may produce generosity as an occurrent passion when one has an idea of something conducive to another person's good (55). Hume did not call the general appetite an instinct, even though he included it among the calm passions. Radcliffe nevertheless takes it to be an innate disposition to seek pleasure and avoid pain and argues that it too counts as an instinct (22–3). Although she does not provide direct textual evidence for this point, this interpretation has some plausibility, since it is an empirical fact that we possess such a disposition innately.

Radcliffe takes this general appetite to play a foundational role in motivation (59–60). She draws attention to the following passage from *Morals*:

> It appears evident, that the ultimate ends of human actions can never, in any case, be accounted for by *reason*, but recommend themselves entirely to the sentiments and affections of mankind, without any dependence on the intellectual faculties. Ask a man, *why he uses exercise*; he will answer, *because he desires to keep his health*. If you then enquire, *why he desires health*, he will readily reply, *because sickness is painful*. If you push your enquiries farther, and desire a reason, *why he hates pain*, it is impossible he can ever give any. This is an ultimate end, and is never referred to any other object. (M App1.18)

Here Hume described hating pain or avoiding it as an "ultimate end" and said that this end recommends itself to the sentiments and affections. This suggests that the general appetite sets what Hume here calls the "ultimate ends" of action.

In addition, Hume wrote in the *Treatise*:

> The impressions, which arise from good and evil most naturally, and with the least preparation are the *direct* passions of desire and aversion, grief and joy, hope and fear, along with volition. The mind by an *original* instinct tends to unite itself with the good, and to avoid the evil, tho' they be conceiv'd merely in idea, and be consider'd as to exist in any future period of time. (T 2.3.9.2)

Draft. Do not Cite.

Read in light of the passage from *Morals* quoted above, the "*original* instinct" mentioned here may be interpreted as referring to the general appetite (Radcliffe 2018, 22–3). If this interpretation is correct, the following psychological process of motivation can be reconstructed. First, when an agent forms a belief about pleasure or pain, for example, the belief that sickness is painful, the general appetite, that is, the disposition to avoid pain, is activated. It then produces an occurrent desire directed at a related particular object, such as the desire to preserve one's health. Finally, together with a belief about the means of preserving one's health, this desire produces a desire for a specific course of action such as exercising.[7] According to Radcliffe, the other innate dispositions produce motivating passions through the same sort of process (55).

On this basis, the core of the instinctive interpretation is the view that the calm passions are innate dispositions, and that, under particular circumstances, they produce further occurrent passions that can serve as motives. This interpretation can avoid both the justificatory difficulty and Qu's objection from triviality. In the case of the general appetite, it is hardly deniable that we are innately disposed to seek pleasure and avoid pain. Moreover, with regard to the other calm passions such as benevolence and resentment, it is at least possible to hold that we possess such tendencies innately, and it also seems possible to adduce empirical support for that claim, although whether it is in fact correct is another matter. According to this interpretation, then, Hume neither inferred the involvement of the calm passions in motivation retrospectively from action nor treated it as a mere conceptual requirement.

Before proceeding, a conceptual clarification is needed concerning instinct, which will be important for the argument that follows. The fact that Hume described something as an instinct does not by itself imply that it is a disposition. Hume used the term "instinct" in a variety of contexts (see, for example, T 3.2.2.16), so its meaning in any particular passage should not be assumed in advance to be that of a disposition. Nor does Hume's use of the term "instinct" by itself show that what he called an instinct is a calm passion. As discussed in Section 1, the distinction between calm and violent concerns passions and the degree of emotion they produce. Whether what Hume called an instinct is itself a passion, however, is a separate question that must be settled on the basis of the text. Radcliffe's interpretation links these three notions by identifying the relevant instincts with innate dispositions that are themselves the calm passions. Whether Hume's texts warrant this combination, however, is a further interpretive question.

In Section 4, I criticise this identification of calm passions with instincts and offer an alternative account of Hume's use of "instinct" in connection with calm passions, namely, as referring to a principle that gives rise to a passion under specific circumstances. For present purposes, this principle may be understood in terms of a cause-effect relation in which a specific circumstance is followed by a response that is original rather than acquired through experience or engineered by artifice. Understanding the principle in this way need not by itself commit us to any abiding psychological feature underlying the response and can therefore serve as an alternative to Radcliffe's dispositional account of instinct.

3. Difficulties for the Instinctive Interpretation

Several criticisms of the instinctive interpretation have been raised. The first is that, because Hume did not explicitly identify the general appetite as an instinct, the interpretation cannot be sustained (Kail 2007, 196–7; Sturgeon 2015, 263–4). But since, as Section 2 has shown, there

[7] Unlike Radcliffe, other supporters of the instinctive interpretation do not explicitly discuss the general appetite. However, they agree with her in taking the calm passions to be innate dispositions that, under certain conditions, give rise to motivating passions.

Draft. Do not Cite.

is at least some textual support for this interpretation, it would be premature to reject it in its entirety on this ground alone.

Rachel Cohon develops a more theoretical criticism by focusing on the vulgar concept of "power" that Hume developed in Book 2 of the *Treatise*. While Hume's strict philosophical definition equates a power with its actual exercise, this vulgar concept allows for unexercised abilities, defining power simply as the probability that a specific effect will follow given the right conditions. Cohon (2008, 56) takes this concept of power to be a disposition whereby one thing produces another under certain conditions. When applied to motivation, this means that the bearer of the power to produce a desire must be some mental state other than the disposition itself. The instinctive interpretation, therefore, which treats the calm passions themselves as possessing this power, is inconsistent with this concept of power (56–9).

However, this criticism has its own difficulties. Although Cohon identifies the vulgar concept of power with disposition in Hume, she does not provide sufficient textual evidence for doing so. Hume himself never discussed the concept of disposition explicitly, and on this point, Cohon says only that "[t]he obvious place to seek his position on dispositional properties is his account of causal powers" (55). Without substantiating this identification, defenders of the instinctive interpretation may read into Hume a concept of disposition that supports their view.

Existing criticisms are thus suggestive, but they do not by themselves get at the core problem with the instinctive interpretation. By contrast, what this paper emphasizes is consistency with Hume's framework of impressions and ideas. As Qu also notes (2018, 445n4), Hume said, in his discussion of the idea of the self, that "there is no impression constant and invariable. Pain and pleasure, grief and joy, passions and sensations succeed each other, and never all exist at the same time" (T 1.4.6.2). The calm passions as innate dispositions, especially the general appetite, seem to come uncomfortably close to being something "constant and invariable." Moreover, the instinctive interpretation seems inconsistent with Hume's remark that "the impressions of reflection, *viz.* passions, desires, and emotions, which principally deserve our attention, arise mostly from ideas" (T 1.1.2.1), since on that interpretation the direct cause of desires would be passions as innate dispositions.[8]

Next, T 2.3.9.2 and M App1.18, which provide the basis for the instinctive interpretation, do not directly establish its dispositional claim. Radcliffe takes the original instinct to seek good and avoid evil mentioned in T 2.3.9.2 to be an ultimate cause of the direct passions. Hume also acknowledges that some direct passions arise from other natural impulses or instincts (T 2.3.9.8). What does not follow, however, is that either kind of instinct must itself be understood as an abiding calm passion or disposition.

For its part, M App1.18 argues only that what Hume calls the "ultimate ends of human actions," such as avoiding pain, "recommend" themselves to sentiment rather than to the intellectual faculties. Hume was not here characterizing this sentiment itself as an instinct or a disposition, nor did he say that it is the cause that produces motivating passions. Thus, deriving the instinctive interpretation from this argument would require reading a good deal into the text. Furthermore, the interpretation is problematic in its handling of the texts. Hume was primarily discussing the regress generated by the question concerning the reason for action, namely, why one desires something, not the psychological causes of action and the passions, as in the *Treatise*. Since the two works discuss action from different points of view, their respective contexts should be carefully distinguished in interpretation.[9] Indeed, even if avoiding pain were

Draft. Do not cite.

[8] Radcliffe (2018, 39–52) resists the view that ideas or beliefs produce most desires, arguing instead that desire arises from impressions in conjunction with the general appetite. More specifically, on her view, an impression of pleasure or pain previously experienced by the agent reappears, is projected by inference onto a future object, and produces a desire in combination with the general appetite. This account, however, remains in strong tension with the natural reading of Hume's claim that the idea of pleasure or pain produces desire and aversion (T 1.1.2.1), together with his insistence that an impression reappears in the mind as an idea (T 1.1.3.1).

[9] Of course, it is possible in principle to interpret the *Treatise* in terms of the concept of a reason for action, as Bricke

the ultimate end of our actions, this would not establish that particular desires are produced by a pre-existing general aversion to pain (Persson 1997, 199–200).

These considerations, then, suggest that the instinctive interpretation faces serious interpretive difficulties. Although they do not by themselves suffice to refute the instinctive interpretation outright, they do require a significant qualification of its explanatory force. I therefore turn next to benevolence, one of the calm passions, in order both to deepen the critique of the instinctive interpretation and to present my own interpretation based on the occurrent state interpretation.

4. Benevolence, Instinct, and the Calm Passions

4.1 Benevolence and Instinct

To clarify at the outset what is at issue between the instinctive interpretation and my own, I reconsider the passions connected with instinct. So far, the focus has been on the calm passions listed by Hume at T 2.3.3.8, namely benevolence, resentment, the love of life, kindness to children, and the general appetite. But Hume also distinguishes direct from indirect passions, a classification independent of the distinction between calm and violent passions (T 2.1.1.3–4). This second classification is relevant here because Hume also discusses the relation between direct passions and instinct in the *Treatise*, particularly in the following passage:

> Beside good and evil, or in other words, pain and pleasure, the direct passions frequently arise from a natural impulse or instinct, which is perfectly unaccountable. Of this kind is the desire of punishment to our enemies, and of happiness to our friends; hunger, lust, and a few other bodily appetites. These passions, properly speaking, produce good and evil, and proceed not from them, like the other affections. (T 2.3.9.8)

Some might appeal to passages like T 2.2.9.3 to claim that "the desire of punishment to our enemies, and of happiness to our friends" mentioned in T 2.3.9.8 refers to the passions of anger and benevolence respectively. On this basis, proponents of the instinctive interpretation could take anger, hunger, and lust, along with the calm passions listed at T 2.3.3.8, to be innate dispositions (Radcliffe 2018, 22).

However, as Radcliffe (2018, 22) notes, the passage from T 2.3.9.8 quoted above admits of two readings. The key question is what "[o]f this kind" refers to. The interpretation differs depending on whether the phrase refers just to "a natural impulse or instinct" or to "the direct passions" that do not arise from good and evil but from an unaccountable natural impulse or instinct. Take benevolence as an example. If the former reading is correct, the passion of benevolence is itself an instinct, and this instinct produces one of the direct passions. This reading provides direct support for the instinctive interpretation. If, by contrast, the latter reading is correct, benevolence is one of the direct passions arising from instinct, rather than instinct itself.

The latter reading seems more natural in the context of T 2.3.9.8. The phrase "[o]f this kind" appears to refer to something that has been divided into kinds. That would be the direct passions, which have been divided into a kind arising from good and evil and another arising from a natural impulse or instinct, the latter being the particular kind referred to in the following sentence. Natural impulses or instincts have not been divided into kinds, and so are less plausibly what "[o]f this kind" refers to. I wish to further defend this latter reading in this section.[10]

(1996) does, for example. However, reasons for action are not explicitly discussed in the *Treatise*, and any interpretation that aims to remain faithful to Hume's text must therefore provide some justification for introducing this concept. Yet, at least in the arguments advanced by the defenders of the instinctive interpretation considered in this paper, such justification has not been adequately provided.

[10] Cohon (2008, 35) also distinguishes instincts from the direct passions they produce, but identifies hunger, lust,

Draft. Do not Cite.

To show this, I begin by considering how benevolence, the only passion that (conjecturally) figures in both T 2.3.3.8 and T 2.3.9.8, arises.[11] As the preceding discussion has already suggested, benevolence is a kind of desire (T 2.2.9.3, T 2.3.3.8). However, unlike ordinary desire which arises from good or evil, the benevolence discussed in T 2.2.6 arises in conjunction with another passion, namely, love (T 2.2.6.3). Hume wrote:

> We may, therefore, infer, that benevolence and anger are passions different from love and hatred, and only conjoin'd with them, by the original constitution of the mind. … According as we are possess'd with love or hatred, the correspondent desire of the happiness or misery of the person, who is the object of these passions, arises in the mind, and varies with each variation of these opposite passions. (T 2.2.6.6)

What is important in this passage is that Hume distinguished love from benevolence. Indeed, before this quotation, Hume wrote that benevolence and anger "arise only upon the ideas of the happiness or misery of our friend or enemy being presented by the imagination" (T 2.2.6.5). The benevolence in question should therefore be understood as an occurrent desire that arises when one has love for a person and the idea of that person's happiness is presented by the imagination.

In the next section, Hume wrote that "the desire of the happiness or misery of others, according to the love or hatred we bear them, [is] an arbitrary and original instinct implanted in our nature" (T 2.2.7.1). Since Hume here apparently took this desire to be instinctive, this passage, taken in isolation, might suggest that he is discussing this desire as an innate disposition. As shown above, however, immediately before this he discussed benevolence as an occurrent passion. In that context, it is therefore more plausible to think that Hume was referring to an instinct to feel an occurrent passion, benevolence, when presented with the right stimulus, an idea of the happiness of someone we love. This point is important for the present argument, because it shows that, although Hume referred to one of the calm passions at T 2.3.3.8 simply as an instinct, this does not by itself support the dispositional or instinctive interpretation.

4.2 Hume's use of "instinct"

It remains unclear, however, what Hume meant when calling an occurrent passion an "instinct" with no further qualification. I therefore examine below Hume's use of the term "instinct" in passages of the *Treatise* and *Morals* that have not figured in the discussion so far and use that examination as a guide for interpreting T 2.3.3.8 and T 2.3.9.8.

I begin by analysing Hume's use of "instinct" in the *Treatise*, but three limitations should be noted. First, although Hume used the term in a variety of contexts in this work, the following analysis is confined, given the concern of this paper, to cases in which the term is used in connection with the passions. Secondly, as noted above, I set aside for the moment those passages that have already been at issue in the preceding discussion, and consider other occurrences first.[12] Thirdly, because Hume sometimes used the expressions "original quality" and "primary constitution" in explanatory contexts closely related to his use of "instinct" (T 2.1.5.3, T 3.1.2.6), these expressions will also be considered where relevant.

An analysis of the passages under these constraints shows that Hume used "instinct" and the related expressions chiefly in two ways: (a) to refer to a principle that determines certain

anger, and benevolence as instincts themselves rather than as direct passions arising from instinct.

[11] Although T 2.3.9.8 does not explicitly mention benevolence, its reference to the desire for the happiness of our friends can plausibly be understood as referring to benevolence in light of T 2.2.6.3–6 and T 2.2.9.3. This identification is not conclusively established by T 2.3.9.8 alone. I shall nonetheless adopt it here for the purpose of engaging with Radcliffe's interpretation (2018, 22).

[12] The passages excluded here are T 2.2.6.6, T 2.2.7.1, T 2.3.3.8, T 2.3.9.2, T 2.3.9.8, and T 3.3.1.31. Although T 3.3.1.31 has not been discussed so far, its content is close to that of T 2.2.6.6, and it is therefore excluded from the present analysis as well.

Draft. Do not Cite.

structural features of a passion (T 2.1.3.5, T 2.1.5.3, T 2.1.8.2); and (b) to refer to a principle that gives rise to a passion under specific circumstances (T 2.1.11.11, T 2.2.12.5, T 3.1.2.6). Typical instances of (a) are found where Hume discussed the object of a passion (T 2.1.5.3) and the adaptation of certain qualities to it (T 2.1.3.5), whereas (b) is illustrated by the passage in which parental affection is said to proceed from instinct (T 2.2.12.5). I shall return to the use of "instinct" described in (b) at the end of this section to clarify how this principle is to be understood.

I now turn to *Morals*. In this work, most occurrences of the term "instinct" appear in the context of Hume's discussion of whether the sentiment of justice is an instinct (M 3.40–3.44, 3.46). At the outset of that discussion, Hume wrote:

> The dilemma seems obvious: As justice evidently tends to promote public utility and to support civil society, the sentiment of justice is either derived from our reflecting on that tendency, or like hunger, thirst, and other appetites, resentment, love of life, attachment to offspring, and other passions, arises from a simple original instinct in the human breast, which nature has implanted for like salutary purposes. If the latter be the case, it follows, that property, which is the object of justice, is also distinguished by a simple, original instinct, and is not ascertained by any argument or reflection. (M 3.40)

This passage refers, albeit in slightly different terms, to the passions discussed in T 2.3.3.8 and T 2.3.9.8. It says that these passions "arise from a simple original instinct," and thus uses "instinct" as described in (b). After this quotation, Hume continued to use the term "instinct" in the context of rejecting the view that property can be distinguished by instinct. It is therefore reasonable to think that he used "instinct" as described in (b) in the discussion that follows as well.[13]

Apart from this context, the word "instinct" is used in *Morals* only at M App3.2 and M App3.9n. The latter is not directly concerned with the passions. In the former, Hume said that "[t]he social virtues of humanity and benevolence exert their influence immediately, by a direct tendency or instinct" (M App3.2). Although this passage alone does not determine the relation between instinct and passions such as benevolence, neither does it mandate identifying the two.

The foregoing completes the analysis of Hume's uses of "instinct" in the *Treatise* and *Morals*. Within this analysis, the discussion of M 3.40 is especially important for the present argument. For it presents resentment, the love of life, and kindness to children (attachment to offspring) as arising from instinct rather than identifying these passions with instinct itself. This alone, however, does not warrant the conclusion that Hume used the term "instinct" in the same way in T 2.3.3.8, where the *Treatise* refers to these passions. At first glance, T 2.3.3.8 appears to identify these passions with instinct. But there is reason to think that Hume distinguished the calm passions from instinct, even in T 2.3.3.8, and took the latter to produce the former.[14] To show this, I shall reconsider Hume's discussion of benevolence in T 2.2.7.1.

Given the foregoing discussion, it is natural to read "instinct" in T 2.2.7.1 as described in (b). As argued in Section 4.1, the benevolence discussed in T 2.2.7.1 is best understood as an

---

[13] Hume's discussion from M 3.41 onward is somewhat ambiguous, and one might read "instinct" there in a way other than those described in (a) or (b). I do not claim, however, that every occurrence of the term in connection with the passions must be read in one of those two uses. For present purposes, it is enough that M 3.40 exhibits the use described in (b). Even if other readings are possible later in the discussion, they do not undermine the interpretation advanced here.

[14] It should be noted that there is no principled reason to exclude the use of passages from *Morals* in interpreting the *Treatise*. My criticism of Radcliffe in Section 3 is directed not at such use itself, but at her use of those passages without sufficiently distinguishing the different contexts of the discussions in the two works. In the present case, Hume refers in M 3.40 to the passions discussed in T 2.3.3.8 and T 2.3.9.8. Although this discussion was added only in later editions of Morals, it is compatible with the relevant passages in the Treatise and may therefore be used as a guide to their interpretation.

Draft. Do not Cite.

occurrent passion. If we do not interpret it in this way, Hume would appear, without any indication, to have shifted from discussing benevolence as an occurrent passion in the immediately preceding paragraph to discussing it as something ontologically different. Moreover, the direct cause of this occurrent benevolence is not good or evil. If it were, Hume would presumably have included benevolence among the passions that arise from good and evil (T 2.3.9.5–7). In light of these considerations, benevolence can be understood as an occurrent passion produced by instinct.[15] The relevant environmental condition, I suggest, is that the imagination presents the idea of something that affects a friend's happiness (T 2.2.6.5). If this reading is correct, then the fact that Hume appears in T 2.2.7.1 to identify benevolence with instinct can be explained as a mere abbreviation of the fuller account, and this use of "instinct" should be understood as described in (b). T 2.2.6.6 supports this reading of T 2.2.7.1, and T 2.3.9.8 supports a similar reading of T 2.3.3.8.

If this interpretation is correct, then, among the passions listed there, all except the general appetite may be understood as passions produced by an instinct other than the instinct to pursue good and avoid evil. The (b) interpretation of "instinct" is grounded in the most natural reading of T 2.3.9.8. Moreover, this reading is supported by M 3.40, where Hume used "instinct" as described in (b) in connection with bodily appetites such as hunger and thirst. These considerations support the central interpretation advanced in this paper: calm passions are, with the exception of the general appetite, best understood as occurrent passions produced by instinct.

The foregoing analysis also allows us to develop the account of instinct anticipated in Section 2. Hume does not explicitly explain the status of instinct in the usage described in (b), as illustrated by his brief remark that parental affection "proceeds from a peculiar instinct" (T 2.2.12.5). Nevertheless, the relevant use of "instinct" can be understood as referring to an original principle—one neither acquired through experience nor engineered by artifice—that gives rise to a passion under specific circumstances. This is not to say (or deny) that there is anything in the agent that abides and makes it react as it does to the circumstances it encounters. But the instinctive reaction can be understood independently of identifying any such abiding feature and can consistently be thought to figure in a brute factual regularity, resistant to further explanation (E 5.5). This reading is also consistent with Hume's broader discussion of "instincts", where he describes operations that occur when the mind is placed in certain circumstances—including feeling love upon receiving benefits and hatred upon meeting with injuries—as "a species of natural instincts," which no reasoning or process of thought and understanding is able either to produce or to prevent (E 5.8). In the cases relevant here, the response is the passion. On this reading, then, Hume's use of "instinct" can be given a positive account without treating the calm passions themselves as abiding factors responsible for the connection between a circumstance and an occurrent passion.

4.3 Textual Advantages

The foregoing discussion has established the central interpretation advanced in this paper. I now turn to the interpretive advantages of this reading, focusing in particular on the ways in which it allows for a more consistent and textually coherent understanding of Hume's discussion.

The chief advantage of this interpretation lies in its ability to explain the relevant texts in the

[15] In the *Treatise*, benevolence has more than one cause. According to Hume, we can feel benevolence even toward persons who are not objects of love, and in such cases, benevolence arises through sympathy (T 2.2.9.13) or through the parallel direction of the affections (T 2.2.9.9). On this point, Vitz (2004, 265) argues that Hume distinguishes between benevolence in a narrow sense, which accompanies love and arises from instinct, and benevolence in a broader sense, which arises through sympathy or the parallel direction of the affections. It should therefore be noted that the interpretation advanced here applies only to benevolence in the narrow sense.

Draft. Do not Cite.

Treatise and *Morals* more coherently than the instinctive interpretation does. As noted above, Hume presented benevolence as an occurrent desire for the happiness of the person one loves (T 2.2.6–7). On my reading, Hume may also be read as discussing occurrent benevolence in the same sense in T 2.3.3.8 and T 2.3.9.8. By contrast, on the dispositional interpretation, including its instinctive variant, Hume would have discussed benevolence as an occurrent passion in T 2.2.6–7, but benevolence as a disposition in T 2.3.3.8 (Radcliffe 2004, 647n28; 2018, 21–3). Hume himself, however, gave no explicit indication of any such equivocal usage.

Moreover, the present reading allows for a more consistent interpretation of Hume's discussion across the *Treatise* and *Morals*. In M 3.40, Hume distinguished passions such as resentment, the love of life, and attachment to offspring from instinct, while also presenting instinct as their source. Read in this light, T 2.3.3.8 is best understood not as identifying these passions with instinct, but as abbreviating a fuller explanation already compatible with Hume's treatment elsewhere.

A further textual advantage is that this reading makes better sense of Hume's claim that certain passions "produce good and evil and proceed not from them" (T 2.3.9.8). Hume seems not to have explained this claim in detail in the *Treatise*. However, in *Morals*, he said that hunger and thirst do not arise from pleasure and pain, but that pleasure arises from their gratification (M App2.12). He made a similar point about benevolence, writing that "from the original frame of our temper, we may feel a desire of another's happiness or good, which, by means of that affection, becomes our own good" (M App.2.13). Read together, these passages suggest that hunger, thirst, and benevolence are not passions arising from good and evil, but passions that themselves produce good and evil. More importantly, Hume described them here as passions that are felt or gratified, so that the present interpretation makes better sense of these passages than dispositional interpretations, including the instinctive variant.

5. Implications for Hume's Theory of Passions and Motivation

This section draws out four theoretical implications for our understanding of Hume's theory of passions and motivation that follow from the present interpretation, while also clarifying its relation to the occurrent state interpretation more generally.

First, on the present interpretation, the passions listed at T 2.3.3.8 (with the exception of the general appetite) and T 2.3.9.8 are to be understood as occurrent passions produced by exceptional processes. The direct passions including desire are in general produced by good and evil (T 2.3.9.5–7), while the calm passions may be understood as arising exceptionally from instinct in response to certain circumstances. It should be added, however, that even if instinct were the ultimate cause of the direct passions, this would not undermine the central claim advanced here. What is rejected here is the notion that instinct itself is a calm passion, not the claim that instinct may play some causal role in motivation. Rather, on the present interpretation, calm passions such as benevolence arise from instinct, and it is those passions that function as motives.

Second, on my reading, both calm and violent passions are to be understood as occurrent mental states. This makes it easier to account for Hume's discussion of the continuity between the two. Near the end of his discussion of the will, Hume reiterated the distinction between calm and violent passions, as well as the tendency of the former to be confused with the operations of reason and then said that calm passions can turn into violent passions as circumstances or objects change (T 2.3.8.13). On a natural reading, the calm passions in this passage can be understood as referring to those in T 2.3.3.8. Indeed, according to Hume, resentment may appear as a violent passion when one has suffered direct injury (T 2.3.3.9). Read together with his claim that "The same good, when near, will cause a violent passion, which, when remote, produces only a calm one" (T 2.3.4.1), it is plausible to think that, at least in the case of a passion such as resentment, one and the same passion may become either calm or violent

Draft. Do not cite.

depending on the circumstances (Baier 1991, 167).

By contrast, Radcliffe (2015, 552–6) understands this continuity to mean that one and the same passion may appear either in a calm mode or in a violent mode. The former is a settled principle of action; the latter is a passion accompanied by strong emotion. On her view, these two modes are continuous in terms of force and vivacity. However, this continuity between calm and violent passions does not fit with a dispositional or instinctive interpretation. Accepting that violent passions are clearly occurrent, if the calm mode signifies a disposition, then the continuity in question is one between dispositional and occurrent states, despite their differing ontological status. If, on the other hand, the calm mode signifies an occurrent state with a low degree of force and vivacity, then the occurrent state interpretation is endorsed.

Defenders of the dispositional interpretation, including the instinctive variant, might reply that Hume himself recognized a continuity between dispositions and passions. Hume did indeed say that a passion may become a settled principle of action and, through repetition and habit, come to operate without agitation (T 2.3.4.1, T 2.3.5.1). Two points should be made, however. Firstly, what Hume claimed in this discussion is that direct passions in general may shift between the calm and the violent in response to changes in circumstances or habit; it is difficult to derive the conclusion that calm passions such as benevolence and resentment are (innate) dispositions. Secondly, this discussion can equally well be read as describing the process by which an occurrent passion comes, through repetition and habituation, to operate more calmly, and Hume himself seems to have suggested as much (T 2.3.5.2).

Third, in connection with the first and second points, the calm passions listed in T 2.3.3.8 can themselves function as motives for action, and when they do so while remaining calm, we mistake them for the operations of reason. Moreover, these passions may at times become intense and still function as motives (T 2.3.3.9). In this respect, the calm passions function in motivation in the same way as ordinary desires.

This point is also important for understanding Hume's account of weakness of will. There Hume said that both calm and violent passions act upon the will, and that, when the two conflict, one or the other prevails (T 2.3.3.10). On the instinctive interpretation, by contrast, the conflict is not with the calm passion itself, but with some distinct occurrent desire produced by it. Yet, if one follows the sequence of T 2.3.3.8–10, it is at least more natural to read Hume as having thought that the calm passions listed in T 2.3.3.8 are themselves in conflict with violent passions, since in T 2.3.3.9 he referred back to "these calm passions, which often determine the will," before claiming in T 2.3.3.10 that both principles operate on the will and may conflict. If so, this passage counts against the instinctive interpretation, even if it does not by itself rule out Stroud's dispositional interpretation.

Fourth, the present interpretation can, to some extent, mitigate the justificatory difficulty that confronts the occurrent state interpretation. In the context of motivation, this difficulty arises because, even if reason alone does not provide a motive for action, this alone does not justify supposing that a calm and occurrent passion is present in every case in which reason appears to do so. This paper shows that at least in the case of some calm passions, such as benevolence, this difficulty does not apply in any straightforward way. For prior to T 2.3.3.8, Hume explained benevolence as an occurrent desire together with the conditions of its genesis and then drew on this passion in explaining cases in which we seem to act in accordance with reason. Moreover, as argued above, the fact that this passion can become extremely calm can be empirically explained within the broader framework of Hume's theory of the passions. Hume's reference to at least benevolence in T 2.3.3.8 may be understood not as an ad hoc hypothesis, but rather as a redeployment, under the heading of motivation, of a passion he had already discussed independently.

This explanation, however, cannot be extended to the other calm passions, such as resentment and the love of life. In his major works, Hume did not provide a detailed account of the conditions under which calm passions other than benevolence arise. As for the general

Draft. Do not Cite.

appetite, not only is no detailed explanation provided, but it is mentioned in only one passage of the *Treatise*. It is not mentioned in *A Dissertation on the Passions*. In that sense, it is impossible to offer a fully text-based explanation of these passions. But at least in the case of benevolence, since Hume did provide a psychological explanation of its genesis, there is no need to regard his discussion of the calm passions as ad hoc.

Conclusion

In this paper, by focusing especially on the conditions under which benevolence arises and on Hume's use of the term "instinct," I have argued that the calm passions, with the exception of the general appetite, should be understood as occurrent passions produced by instinct, not as instincts in their own right. In developing this interpretation, I have clarified the relation between the calm passions and instinct, while also specifying the conditions under which benevolence arises. Owing to the paucity of Hume's own discussion, it has not been possible to specify the conditions under which resentment, the love of life, and kindness to children arise. Nevertheless, this paper has at least shown that these too should be understood as passions produced by instinct.

This interpretation sheds new light on the debate over the ontological status of calm passions by resituating the occurrent state interpretation within Hume's overall system and thereby reinforcing it on a more textually grounded basis. Existing versions of this interpretation have not sufficiently clarified either the conditions under which the calm passions arise or their relation to instinct. By contrast, the dispositional interpretation, and especially its instinctive variant, accounts for calm motivation by treating the calm passions themselves as standing dispositions, but this identification sits uneasily with Hume's psychological framework. The present interpretation helps to address difficulties on both sides and offers a more convincing account of Hume's calm passions.

In a broader context, this interpretation has significant implications for debates over Hume's theory of motivation. While a prevailing view in recent Hume scholarship is that beliefs about pleasure or pain produce desires, interpretations that identify calm passions with instincts have served as a formidable rival to that view (Cohon 2008, 30–62; Radcliffe 2018, 29–64). Accordingly, if my interpretation is correct, it calls for some revision of the current landscape of this debate, since it brings out major difficulties with the instinctive interpretation. Admittedly, this paper does not resolve the broader issues with Hume's theory of motivation, such as how ordinary desires arise from ideas or what relation they bear to reason. Nevertheless, it offers an important step toward understanding his account of motivation in a manner consistent with his own texts.

Acknowledgments

I would like to thank the anonymous reviewers for their careful and constructive comments on an earlier version of this paper. I am especially grateful to the editors, Lorne Falkenstein and Juan Samuel Santos Castro, for the exceptional care and attention they devoted to the manuscript and for their thoughtful and detailed guidance throughout the revision process. Their suggestions prompted me to reconsider important aspects of my argument and greatly helped me clarify and strengthen the paper, particularly my account of instinct and its relation to the calm passions. I deeply appreciate the time and effort they invested in helping me improve this paper.

Draft. Do not Cite.

References

Árdal, Páll S. 1966. *Passion and Value in Hume's Treatise*. Edinburgh University Press.

Baier, Annette C. 1991. *A Progress of Sentiments: Reflections on Hume's Treatise*. Harvard University Press. https://doi.org/10.4159/9780674020382.

Bricke, John. 1996. *Mind and Morality: An Examination of Hume's Moral Psychology*. Clarendon Press. https://doi.org/10.1093/acprof:oso/9780198250111.001.0001.

Cohon, Rachel. 2008. *Hume's Morality: Feeling and Fabrication*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268443.001.0001.

Enç, Berent. 1996. "Hume's Unreasonable Desires." *History of Philosophy Quarterly* 13(2): 239–54.

Fieser, James. 1992. "Hume's Classification of the Passions and Its Precursors." *Hume Studies* 18(1): 1–17. https://doi.org/10.1353/hms.2011.0417.

Immerwahr, John. 1992. "Hume on Tranquillizing the Passions." *Hume Studies* 18(2): 293–314. https://doi.org/10.1353/hms.2011.0393.

Kail, Peter J. E. 2007. *Projection and Realism in Hume's Philosophy*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199229505.001.0001.

Loeb, Louis E. 2002. *Stability and Justification in Hume's Treatise*. Oxford University Press. https://doi.org/10.1093/0195146581.001.0001.

Persson, Ingmar. 1997. "Hume—Not a 'Humean' about Motivation." *History of Philosophy Quarterly* 14(2): 189–206.

Qu, Hsueh. 2018. "Hume's (Ad Hoc?) Appeal to the Calm Passions." *Archiv für Geschichte der Philosophie* 100(4): 444–69. https://doi.org/10.1515/agph-2018-4003.

Radcliffe, Elizabeth S. 2004. "Love and Benevolence in Hutcheson's and Hume's Theories of the Passions." *British Journal for the History of Philosophy* 12(4): 631–53. https://doi.org/10.1080/0960878042000279305.

Radcliffe, Elizabeth S. 2015. "Strength of Mind and the Calm and Violent Passions." *Res Philosophica* 92(3): 547–67. https://doi.org/10.11612/resphil.2015.92.3.1.

Radcliffe, Elizabeth S. 2018. *Hume, Passion, and Action*. Oxford University Press. https://doi.org/10.1093/oso/9780199573295.001.0001.

Shaw, Daniel. 1998. *Reason and Feeling in Hume's Action Theory and Moral Philosophy: Hume's Reasonable Passion*. Edwin Mellen Press.

Stroud, Barry. 1977. *Hume*. Routledge and Kegan Paul.

Sturgeon, Nicholas L. 2015. "Hume on Reason and Passion." In *The Cambridge Companion to Hume's Treatise*, edited by Donald C. Ainslie and Annemarie Butler. Cambridge University Press. https://doi.org/10.1017/cco9781139016100.013.

Vitz, Rico. 2004. "Sympathy and Benevolence in Hume's Moral Psychology." *Journal of the History of Philosophy* 42(3): 261–75. https://doi.org/10.1353/hph.2004.0059.

Draft. Do not Cite.