

The Fiction of “Moral Sentiment”: The primacy of language in Hume’s moral philosophy

SHINYA AIMATSU

Abstract:

David Hume is one of the classical moral sentimentalists, famously claiming that moral judgments are constituted by a certain kind of feeling, or moral sentiment. In this paper, however, I argue that Hume’s moral theory ultimately separates moral judgment from sentiment and, consequently, places the nature of moral judgment in language. To begin with, keeping in mind the so-called Moral Attitude Problem, I show that it is unclear until the end what kind of sentiment Hume’s “moral sentiment” is. Then, I will examine what is really connected with moral judgment in Hume’s theory in the following steps. First, it is confirmed that some moral judgments based on the “general point of view” that Hume introduces are merely verbal judgments without any perceptible sentiments. However, since it is possible that moral language—a set of moral terms—may shoulder the sentimental aspect of such verbal judgments, I trace back the genealogy of moral judgments to see what kind of sentiment was tied to moral language. It is found that when people repeatedly experience that agreement in their reactions to a person is achieved by ignoring the feelings of their sentiments, they establish the general point of view, based on which moral language is created. This origin of moral language suggests that it cannot have any emotional meaning. Thus, insofar as some of what Hume counts as moral judgments do not involve substantive sentiment in any sense, sentiment does not determine the extension of moral judgment. When Hume nevertheless says that moral judgment is constituted by sentiment, that sentiment is merely a kind of fiction, retrospectively assumed from the use of moral language. Finally, given the above, I further argue that it is language that essentially determines the extension of moral judgment in Hume’s moral philosophy.

Keywords:

David Hume, Moral sentimentalism, Moral attitude problem, General point of view, Moral language

1 Introduction

It is a common experience that moral judgments we make are accompanied by some sentiments or emotional attitudes. For example, if we hear that a small child has been abused, we usually judge, with intense discomfort and anger, that such a thing is unforgivable. However, then, whether moral judgment itself is intrinsically tied to sentiment—in particular, whether to make moral judgment is nothing but to have a kind of sentiment (rather than belief, or representative mental state) and whether the function of moral statements is to express a kind of sentiment (rather than belief)—is not easy to answer and is indeed one of the fundamental questions in the history of ethics.

In this paper, I aim to shed some light on this issue through a new interpretation of David Hume’s moral philosophy.¹ Hume proposed a moral theory that can be said to be one of the origins of so-called “moral sentimentalism,” and he believed that moral judgment is essentially constituted by sentiments. “To have the sense of virtue, is nothing but to *feel* a satisfaction of a particular kind from the contemplation of a character. The very *feeling* constitutes our praise or admiration” (T 3.1.2.3). Hume’s moral philosophy has profoundly influenced many contemporary ethicists, inspiring a wide variety of sentimentalist theories (e.g., Blackburn 1993; Slote 2010).

My aim here, however, is not to defend or develop moral sentimentalism based on Hume’s theory, but rather to deny it, at least in part. The interpretation that this paper offers is that in Hume’s moral philosophy—at least in its final form—perhaps contrary to his own awareness, moral judgment is not essentially linked to sentiment, but rather to language, and that the “moral sentiment” that Hume considers to be the nature of moral judgment can be nothing more than a theoretically required fiction, without any definite substance.

The path to this conclusion is as follows. Firstly, in fact, it is not clear what kind of sentiment Hume considered moral sentiment, so much so that this has become one of the traditional problems of interpretation. I suspect that this unclarity stems from the fundamental

¹ All citations to Hume’s *Treatise of Human Nature* (Norton edition, Oxford UP, 2007) will be designated T book.part.section.paragraph; to his *Enquiry Concerning the Principles of Morals* (Beauchamp edition, Oxford UP, 1998) will be designated EPM section.paragraph. Also, note that in this paper, all underlines in quotations are mine.

fictionality of Hume's "moral sentiment," but first, to show how difficult it is to resolve this problem, I will, keeping in mind the so-called Moral Attitude Problem, examine Hume's official answers and an interpretative answer based on the "general point of view" he introduces and show that they all fail (Section 2). Secondly, noting that Hume's general point of view leads to the dilution of sentimental aspects of moral judgment and the emphasis on the role of language in making moral judgments, I argue that at least some of the "moral sentiments" to which Hume refers are merely assumed to exist from the application of moral terms and that, therefore, language determines the extension of moral judgment at least in part (Section 3). However, since it remains possible that moral language—a set of moral terms—itself expresses a certain kind of perceptible sentiment as its function, I focus on Hume's genealogical analysis of moral judgments, tracing it back to the creation of moral language and the establishment of the general point of view, and thereby show that moral language is rather linked to faint "sentiments" without any feeling or liveliness (Section 4). Finally, I argue that, for Hume, moral judgments are verbal judgments using moral language in the first place, and that since there is no substantive sentiment of the kind that constitutes every and only such judgment, it is language, not sentiment, that essentially determines the extension of moral judgment in Hume's moral philosophy (Section 5).

I take my inspiration for this heretical interpretation from King (1976).² According to King, Hume in the third book of his *Treatise* tried to causally explain the production of moral sentiments, or perceptions that correspond to moral judgments, based on the mechanism of perceptions in human nature, but failed to explain how sentiments become so general and sharable as to constitute moral judgments expressed by public language; in contrast, Hume in his *Enquiry*, abandoning a causal explanation of moral judgments, started from analyzing moral judgments expressed by language and thereby succeeded in specifying the criteria for such moral judgments and persuasively inferring the principles of human nature necessary for the adoption of such criteria. Based on the above comparison, King concludes that "the *Enquiry* is superior to *Treatise* III both for generating philosophical conclusions relative to its explanatory focus in human nature, and for its approach to morality" (King 1976: 358).

While I agree with most of King's points, I believe that through a genealogical analysis of the "general point of view" (which I think is the intersection between the approach in the *Treatise* from human nature to language and that in the *Enquiry* from language to human

² As an example of the standard interpretation, Garrett (1997: 196) says, "Making such moral distinctions and evaluations does not depend on the possession of language—at least, not on the possession of moral language. In order to feel moral sentiments, and hence to make moral evaluations in this sense, it is not necessary to command any such terms as 'virtue', 'vice', or 'ought'." Perhaps almost all Hume scholars agree with him on this point.

nature), it is possible to depict the process by which natural sentiments become general and sharable enough to constitute moral judgments. It is, however, also a process by which sentiments become diluted, imperceptible, and fictional. Therefore, although sentiments are necessary for the genealogical origin of moral judgments, the “moral sentiments” that are supposed to underpin individual moral judgments must be fictional. This paper will show this and, hopefully, encourage a reconsideration of the connection between sentiments and morality by presenting a model case in which moral judgment itself can, or even should, be understood independently of sentiments, regardless of how much sentiments are involved in the origins of morality or how seemingly connected some moral judgments and sentiments may be.

2. What is “Moral Sentiment” for Hume?

2.1 Moral Attitude Problem and Hume’s Moral Sentimentalism

One of the major obstacles to moral sentimentalism, which holds that moral judgment is essentially linked to sentiment, is the Moral Attitude Problem formulated by Miller (2013: 39).³ In a nutshell, those who claim that the essence (in some sense) of moral judgment is sentiment must answer the question, “What kind of sentiments are moral?”, or in other words, must specify the kind of sentiments that can adequately determine the extension of moral judgment, which is, however, extremely difficult. If it is impossible to specify it, then sentiment is, at best, only incidentally or secondarily involved in moral judgments. Let me clarify this issue by contrasting moral cognitivism with non-cognitivism. If, as cognitivists say, the nature of moral judgment is to represent or describe a recognized moral object (moral fact or property), then whether a judgment is a moral judgment depends on what kind of fact or property it is about. On the other hand, if, as non-cognitivists say, moral judgments do not involve the cognition of moral objects (there is no such thing in the first place) but express some kind of emotional attitude of the judge toward non-moral objects, then whether a judgment is a moral one depends on what kind of attitude it expresses. But if the latter is the case, then what kind of attitude does moral judgment express in the first place? It is important to note here that one cannot rely on the characteristics of the objects of moral judgments. If non-cognitivists did so, they would be admitting that the object itself has “moral” features, and thus surrendering to cognitivists (McDowell 1998: 158; Kauppinen 2010: 225–226). It now becomes extremely difficult to specify the kind of attitude that is common to all and only moral judgments.

³ See also Miller (1998: 110–112), Smith (2001: 107–114), Merli (2008: 30–43), Kauppinen (2010: 226–230), Köhler (2013: 480–487), and Björnsson and McPherson (2014: 3–12).

Miller (2013: 40–42) suggests that there can be two kinds of answers to the Moral Attitude Problem. One is that there is a *sui generis*, specifically “moral” sentiment that is not reducible to other common sentiments. However, this is not informative at all in the sense that it does not explain what moral sentiments are like, and furthermore, there does not seem to be any sentiment that is introspectively or phenomenologically “moral” and that is always connected with moral judgment. The other is to define moral sentiment as any of the other usual sentiments or a composite of them. In this case, however, defining moral judgment in terms of such sentiments seems to inevitably lead to a wider or narrower extension of moral judgment than the range of judgments that we intuitively classify as moral judgment (Merli 2008; Björnsson and McPherson 2014). In fact, the emotional attitudes we have when making moral judgments are so varied that it is difficult to decide that a certain type of sentiment is the moral sentiment that constitutes moral judgment.

Thus, if we place the essence of moral judgments in sentiments, the challenge is to specify the kind of sentiments that can validly determine the extension of moral judgment. This circumstance is undoubtedly true of Hume’s moral theory, which holds that moral judgment is constituted by sentiment and that sentiment “contains not any representative quality” (T 2.3.3.5). In fact, Hume himself recognizes that this problem arises in his moral system, and he gives his own answers. Before that, Hume’s moral sentimentalism in his *Treatise* begins as follows.

[I]t must be by means of some impression or sentiment they [= vice and virtue] occasion, that we are able to mark the difference betwixt them. [...] Morality, therefore, is more properly felt than judg’d of; (T 3.1.2.1)

The next question is, of what nature are these impressions, and after what manner do they operate upon us? Here we cannot remain long in suspense, but must pronounce the impression arising from virtue, to be agreeable, and that proceeding from vice to be uneasy. (T 3.1.2.2)

Now since the distinguishing impressions, by which moral good or evil is known, are nothing but *particular pains* or pleasures; [...] To have the sense of virtue, is nothing but to *feel* a satisfaction of a particular kind from the contemplation of a character. The very *feeling* constitutes our praise or admiration. [...] [I]n feeling that it pleases after such a particular manner, we in effect feel that it is virtuous. (T 3.1.2.3)

This is very classic sentimentalism, which states that to make a moral judgment is essentially to have a certain sentiment (a pleasant or painful, positive or negative feeling). Then, in the last quotation, the expression “*particular* pains or pleasures” suggests that not all pleasures and pains can be linked to moral judgment, and in the next paragraph, an assumed counterargument is presented.

Now it may, in like manner, be objected to the present [= Hume’s own] system, that if virtue and vice be determin’d by pleasure and pain, these qualities must, in every case, arise from the sensations; and consequently any object, whether animate or inanimate, rational or irrational, might become morally good or evil, provided it can excite a satisfaction or uneasiness. (T 3.1.2.4)

If the nature of moral judgment were simply pleasure and pain, then the extension of moral judgment would become overly broad, since things other than the usual objects of moral judgments (such as inanimate objects) can also cause pleasure and pain in humans. Hume is clearly aware of a kind of Moral Attitude Problem here.

And after this, Hume responds to this objection (dealt with in the next subsection), but nevertheless, in the end, what Hume means by “moral sentiment” remains so unclear as to become a traditional interpretive issue. To see for ourselves the difficulty of this issue, let us look at Carlson’s classification of existing interpretations.

1. Moral sentiments are neither direct nor indirect passions, but calm emotions.
2. Moral sentiments are calm direct passions.
3. Moral sentiments are calm versions of the indirect passions of love or hatred.
4. Moral sentiments are unique species of calm indirect passions.
5. Moral sentiments are indirect secondary impressions.

(Carlson 2014: 75)

I will not go into the details, but with so many possible interpretations still pending, it should be clear that Hume’s moral sentiment is textually ambiguous. This in turn means that Hume’s moral sentimentalism is still subject to the Moral Attitude Problem and its resolution is still required.

2.2 Hume's Official Answers: Indirect Passions and Distinctive Feeling

Let us now examine Hume's official answers to the Moral Attitude Problem. They are of two kinds: one is a phenomenological answer that appeals to the distinctive feeling of moral sentiments, and the other is an answer that appeals to the co-occurrence of sentiments toward persons, or "indirect passions" in Hume's terminology. I will begin by examining the second answer.

Secondly, We may call to remembrance the preceding system of the passions, in order to remark a still more considerable difference among our pains and pleasures. Pride and humility, love and hatred are excited, when there is any thing presented to us, that both bears a relation to the object of the passion, and produces a separate sensation related to the sensation of the passion. Now virtue and vice are attended with these circumstances. They must necessarily be plac'd either in ourselves or others, and excite either pleasure or uneasiness; and therefore must give rise to one of these four passions; which clearly distinguishes them from the pleasure and pain arising from inanimate objects, that often bear no relation to us: And this is, perhaps, the most considerable effect that virtue and vice have upon the human mind. (T 3.1.2.5)

Virtue and vice cause not only particular pleasures and pains but also indirect passions, such as pride/love or hatred/humility toward the person to whom virtue or vice belongs. Inanimate objects and other non-personal objects do not cause this kind of passion and thus are not subject to attribution of virtue or vice, that is, to moral judgment.

Leaving aside the detailed interpretive issues, this answer is insufficient to solve the Moral Attitude Problem. Even if we accept that moral sentiments are always accompanied by indirect passions (although it seems to me that they are often not), this will exclude only judgments about non-personal objects from the extension of moral judgment, and it does not limit judgments about persons in the slightest. In fact, what Hume counts as the cause of indirect passions is anything that is in some sense connected to a person and that is itself pleasurable or painful, not limited to virtue or vice (T 2.1.2.5; T 2.2.1.4). And indeed, Hume states that a person's "body" or "fortune" can produce pleasures or pains *on the same principle as that of moral sentiments*.

The pain or pleasure, which arises from the general survey or view of any action or quality of the *mind*, constitutes its vice or virtue, and gives rise to our approbation or blame, which is nothing but a fainter and more imperceptible love or hatred. [...] [T]he

advantages or disadvantages of the *body* and of *fortune*, produce a pain or pleasure from the very same principles. (T 3.3.5.1)

However, it seems that Hume does not consider such pleasures and pains to be moral judgment, and he tries to exclude them from the extension of moral judgment by introducing a different kind of distinction.

All the sentiments of approbation, which attend any particular species of objects, have a great resemblance to each other, tho' deriv'd from different sources; and, on the other hand, those sentiments, when directed to different objects, are different to the feeling, tho' deriv'd from the same source. [...] [W]hen we survey the actions and characters of men, without any particular interest in them, the pleasure, or pain, which arises from the survey (with some minute differences) is, in the main, of the same kind, tho' perhaps there be a great diversity in the causes, from which it is deriv'd. On the other hand, a convenient house, and a virtuous character, cause not the same feeling of approbation; (T 3.3.5.6)

So, Hume believes that the “feeling” of moral sentiments is uniform and unique enough to form an independent category of sentiments and that this difference in feeling can distinguish approbation by fortune from moral judgment. This phenomenological distinction is Hume's first answer to the Moral Attitude Problem, which we will now examine in detail.

Hume's phenomenological answer begins as follows.

[It is] evident, that under the term *pleasure*, we comprehend sensations, which are very different from each other, and which have only such a distant resemblance, as is requisite to make them be express'd by the same abstract term. ... [A]n inanimate object, and the character or sentiments of any person may, both of them, give satisfaction; but as the satisfaction is different, this keeps our sentiments concerning them from being confounded, and makes us ascribe virtue to the one, and not to the other. (T 3.1.2.4)

We attribute virtues and vices only to persons and not to non-personal objects because the feeling of sentiments excited by the former is unique enough to distinguish those sentiments from the others. Clearly, Hume here believes that the appeal to the phenomenological uniqueness of moral sentiments can successfully determine the extension of moral judgment.

Hume goes on to say that even among sentiments toward persons, moral sentiment retains its uniqueness and is distinguished from sentiments based on self-interest, which is another cause of love or hatred (T 2.2.3).

Nor is every sentiment of pleasure or pain, which arises from characters and actions, of that peculiar kind, which makes us praise or condemn. The good qualities of an enemy are hurtful to us; but may still command our esteem and respect. 'Tis only when a character is considered in general, without reference to our particular interest, that it causes such a feeling or sentiment, as denominates it morally good or evil. 'Tis true, those sentiments, from interest and morals, are apt to be confounded, and naturally run into one another. It seldom happens, that we do not think an enemy vicious, and can distinguish betwixt his opposition to our interest and real villainy or baseness. But this hinders not, but that the sentiments are, in themselves, distinct; and a man of temper and judgment may preserve himself from these illusions. (T 3.1.2.4)

Note, however, that Hume's tone here is very modest. He says that we routinely confuse interested sentiments with moral sentiments and that we usually regard those whose interests are merely in conflict with our own as "vicious" based on our interested sentiments (cf. King 1976: 351).⁴ In addition, since Hume describes this kind of confusion as arising naturally and involuntarily through the psychological mechanism of human nature rather than being made consciously or intentionally (T 2.2.3.2, 7–10), he seems to think that ordinary people *cannot distinguish* between moral and interested sentiments. Does this not imply that the feeling of moral sentiments is not so phenomenologically unique, at least when contrasted with that of interested sentiments?

Certainly, Hume says that "a man of temper and judgment" can distinguish between self-interested and moral sentiments, but then does such a person distinguish the two kinds of sentiments *based on the phenomenological difference* of their feelings? If Hume thinks so, it does not seem empirically plausible at all, and in fact, Hume himself later gives a distinction that is not phenomenological at all.

⁴ Hume also says, far more baldly, "When our own nation is at war with any other, we detest them under the character of cruel, perfidious, unjust and violent: But always esteem ourselves and allies equitable, moderate, and merciful. If the general of our enemies be successful, 'tis with difficulty we allow him the figure and character of a man. [...] But if the success be on our side, our commander has all the opposite good qualities, and is a pattern of virtue, as well as of courage and conduct. [...] In short, every one of his faults we either endeavour to extenuate, or dignify it with the name of that virtue, which approaches it. 'Tis evident the same method of thinking runs thro' common life." (T 2.2.3.2)

[W]e over-look our own interest in those general judgments; and blame not a man for opposing us in any of our pretensions, when his own interest is particularly concern'd. We make allowance for a certain degree of selfishness in men; because we know it to be inseparable from human nature, and inherent in our frame and constitution. By this reflection we correct those sentiments of blame, which so naturally arise upon any opposition. (T 3.3.1.17)

This is where Hume, after introducing “sympathy” as a psychological mechanism necessary to produce moral sentiments, further introduces and explains the “general point of view” in order to compensate for the defects of sympathy. What is important now is that the “reflection” mentioned here seems to be a highly intellectual activity of recognizing the inseparability of human nature and selfishness, and therefore tolerating to a certain degree the selfish behavior of others and regarding it as not morally blamable even if it causes disadvantage or displeasure to oneself. This cannot be a mere phenomenological classification of unpleasant feelings.⁵

Because of these circumstances, some researchers interpret Hume’s moral sentiment ultimately as the kind of sentiment produced and regulated by the general point of view, and that this sentiment constitutes moral judgment (Stewart 1976; Brown 2001). Furthermore, some, such as Kauppinen (2010), attempt to solve the Moral Attitude Problem itself based on Hume’s general point of view as understood in this way.⁶ So, let us now examine the specification of moral sentiment based on the general point of view.

2.3 An Etiological Answer: Hume’s “General Point of View”

In the phenomenological answer quoted earlier, Hume states that “’Tis only when a character is considered in general, without reference to our particular interest, that it causes such a feeling or sentiment, as denominates it morally good or evil” (T 3.1.2.4), which seems to suggest the third answer to the Moral Attitude Problem, that is, the etiological specification of moral sentiments (cf. Kauppinen 2010). Furthermore, Hume, thinking that the nature of moral sentiment has already been specified, then asks, “Why any action or sentiment upon the

⁵ Radcliffe says that “we can detach the point about the phenomenon of feeling from that about the general consideration of a character. [...] [This latter point] is central to his sentimentalism in the *Treatise*” (Radcliffe 2022: 17). I agree with her on this point, but, as we shall see, it is this detachability that ultimately detaches Hume’s theory of moral judgment from sentiment.

⁶ However, Björnsson and McPherson (2014: 11) criticize Kauppinen’s criterion of moral sentiment as “over-discrimination,” that is, making the extension of moral judgment narrower than it actually is.

general view or survey, gives a certain satisfaction or uneasiness” (T 3.1.2.11), and spends the rest of his *Treatise* trying to elucidate the mechanism of this causal relation.

However, this etiological characterization is still ambiguous, especially as to what is meant by “considered in general” and “upon the general view or survey.” And indeed, one interpretive controversy has arisen here, namely, whether these phrases imply the “general point of view” introduced much later. With this in mind, let us first overview the mechanism Hume discovered for generating moral sentiments.

According to Hume, the ultimate object of each moral judgment is a person’s mentality, including motive, character, and mental quality (T 3.2.1.2; T 3.3.1.4). When an agent’s character is observed or imagined (through her actions) to bring pleasure or pain (or its cause) to the agent herself or others (T 3.2.2.24; T 3.3.3.3; T 3.3.1.24; T 3.3.4.8), the observer sometimes sympathizes with the pleasure or pain—feels it as if it were his own—and consequently morally approves or disapproves of the agent that caused it (T 3.2.2.24; T 3.3.1.11). Here, sympathy, which is the mechanism that converts others’ actual or imaginable sentiment into one’s own actual sentiment, functions to extend one’s concern from one’s own interests and pleasures/pains to others’ (T 3.3.1.11).

However, Hume then suggests that sympathy alone is insufficient as a principle of moral judgment by bringing up the following supposed objection.

But as this sympathy is very variable, it may be thought, that our sentiments of morals must admit of all the same variations. We sympathize more with persons contiguous to us, than with persons remote from us: With our acquaintance, than with strangers: With our countrymen, than with foreigners. But notwithstanding this variation of our sympathy, we give the same approbation to the same moral qualities in *China* as in *England*. They appear equally virtuous, and recommend themselves equally to the esteem of a judicious spectator. The sympathy varies without a variation in our esteem. Our esteem, therefore, proceeds not from sympathy. (T 3.3.1.14)

The closer the relationship between one person and another, the more strongly and easily sympathy works, and as a result, the stronger sentiments are likely to be evoked (T 2.1.11.4–6). Thus, for example, when you are in England, you are more likely to sympathize more strongly with the English than with the Chinese. Now, if moral sentiment arose directly from sympathy, then when you observe two persons of similar character, one in England and the other in China, you would be more likely to sympathize more strongly with the pleasant or unpleasant influence of the former on the English than that of the latter on the Chinese, and consequently

more strongly morally approve or disapprove of the former. But this is contrary to our practice of moral approbation or disapprobation.⁷

Hume attempts to solve this problem by introducing “some *steady* and *general* points of view” (T 3.3.1.15) or, in other words, “some common point of view” (T 3.3.1.30). For now, let me just outline this point of view itself: That X observes Y’s character from the general point of view means that X sympathizes with the pleasures and pains that Y’s character brings to Y and the people around Y no matter how remote they may be to X. This is contrasted with X’s concern about the pleasures and pains that Y’s character brings to X himself (self-interested point of view) and X’s sympathy with the pleasures and pains that Y’s character brings to people with whom X naturally and easily sympathizes, typically, the people around X (merely sympathetic point of view), both of which together Hume calls X’s “peculiar point of view” (T 3.3.1.15). Hume says,

Being thus loosen’d from our first station [i.e., from our peculiar point of view], we cannot afterwards fix ourselves so commodiously by any means as by a sympathy with those, who have any commerce with the person we consider. (T 3.3.1.18)

[I]n judging of characters, the only interest or pleasure, which appears the same to every spectator, is that of the person himself, whose character is examin’d; or that of persons, who have a connexion with him. [...] They [= such interests and pleasures] alone produce that particular feeling or sentiment, on which moral distinctions depend. (T 3.3.1.30)

If such a special kind of sympathy were available to us, then surely, based on it, observers would be able to perceive the pleasures and pains caused by the character of the person in question in a stable way, whether that person is in England or China. Moreover, such pleasures and pains would appear the same to all observers because those pleasures and pains are fixed around the person in question, in contrast to the observer-relative pleasures and pains that each observer naturally sympathizes with. And Hume says that only such pleasure or pain gives rise to moral sentiment, and thus to moral judgment. This seems to be Hume’s final answer to the Moral Attitude Problem.

⁷ But is this really contrary to our *practice*? As a matter of fact, this kind of bias seems to be common in everyday life. In my view, the assertion that “we give the same approval to the same moral qualities in *China* as in *England*” is really referring to our *conception* of moral judgment, not to our practice of it. As we will see in Section 5, this conception is built into the *meaning of moral language*.

However, despite Hume's conclusion above, some researchers have persuasively argued that taking up the general point of view is not essential to having moral sentiment (Garrett 1997: 196–197, 2001a: 188–189; Cohon 2008: 142–143). One reason for this is that Hume seems to recognize as moral sentiment the observer-relative, variable pleasures and pains that natural sympathy conveys before taking up the general point of view (Garrett 2001b: 211–212; Cohon 2008: 141).

The approbation of moral qualities most certainly [...] proceeds entirely from a moral taste, and from certain sentiments of pleasure or disgust, which arise upon the contemplation and view of particular qualities or characters. Now 'tis evident, that those sentiments, whence-ever they are deriv'd, must vary according to the distance or contiguity of the objects; nor can I feel the same lively pleasure from the virtues of a person, who liv'd in *Greece* two thousand years ago, that I feel from the virtues of a familiar friend and acquaintance. (T 3.3.1.15)

In general, all sentiments of blame or praise are variable, according to our situation of nearness or remoteness, with regard to the person blam'd or prais'd, and according to the present disposition of our mind. (T 3.3.1.16)

These fluctuating sentiments do not seem to be based on the general point of view. Nevertheless, Hume appears to describe them as moral sentiments. Furthermore, immediately after explaining how “we correct those sentiments of blame” based on the general point of view, Hume says,

But however the general principle of our blame or praise may be corrected by those other principles, 'tis certain, they are not altogether efficacious, nor do our passions often correspond entirely to the present theory. (T 3.3.1.18)

If “the general principle of our blame or praise” meant the general principle of *moral* sentiment or judgment, then prior to the intervention of any correction based on the general point of view, moral sentiments themselves would already have arisen from that principle.

There is room to argue against this interpretation, saying that what is fluctuating is “non-moral” sentiments of blame or praise and that it is only by taking up the general point of view that we can have “moral” sentiments of blame or praise. In fact, as Brown points out, since there can be non-moral blame and praise, we need something to distinguish them from moral blame and praise, which, in Hume's theory, seems to be none other than the general point of view as the cause of the latter alone (Brown 2001: 200).

However, Garrett says that for Hume, the distinction between moral and non-moral sentiments is ultimately only a difference in phenomenological feeling, and that, in any case, a dilemma arises for those, like Brown, who hold that the general point of view is necessary for moral sentiments.

Either the moral sentiments are, or they are not, distinguishable by their feeling alone. (1) If Brown holds that they are not, then her view entails that we must first investigate the causal origin of a particular sentiment in order to determine whether it *is* a moral evaluation or not. But Hume asserts that “to have the sense of virtue, is nothing but to *feel* a satisfaction **of a particular kind** from the contemplation of a character. The very *feeling* constitutes our praise or admiration. **We go no farther; nor do we enquire into the cause of the satisfaction**” [T 3.1.2.3]. (2) On the other hand, if Brown grants that Humean moral sentiments *can* be distinguished simply by their feeling, then her claim that they cannot occur without adopting a “general point of view” that involves “regulating our sympathetic responses” is simply a claim about the causal preconditions for the occurrence of sentiments with this particular feeling. But what Brown calls “the general point of view” involves several different kinds of corrections and regulations. Must we perform all of them before we can have moral sentiments? (Garrett 2001b: 212–213; numbers in parentheses are mine)

Regarding the first horn (1), the quote from Hume is probably a description of the way we make moral judgments in our daily lives (while Hume himself is “philosophically” enquiring into the very cause of moral sentiments). It means that we do not make moral judgments after identifying the cause of a particular sentiment we have, but rather we have already made moral judgments at the time we feel that sentiment. Garrett argues that if moral sentiments were not specified by their feeling but by the adoption of the general point of view preceding them, then it follows that, contrary to Hume’s assertion, in order to determine whether or not a sentiment we have is moral and thus constitutes moral judgment, we would have to ascertain whether the sentiment has been caused by our taking up the general point of view. I think this point is compelling not only as an interpretation of Hume but also as a description of our practice. For example, if I observe someone’s behavior from only one aspect and unreflectively judge that he is vicious, no amount of searching for the cause of that judgment will change the fact that I made a moral judgment. Of course, as a result of such a search for the cause, I may find that it was an “unjust” moral judgment, but this will not change the moral judgment that I have already made into a non-moral judgment. Insofar as making a moral judgment is nothing but having a certain sentiment, this story applies to Hume’s theory as well.

Turning now to the second horn (2), I think that this problem arises at least partly because Brown has burdened the general point of view with corrective and regulatory functions that have *degrees* and *no definite end* (Brown 2001: 198). I will not go into the details,⁸ but Garrett's point is that it is strange that one cannot have any moral sentiment and thus cannot make any moral judgments until one has gone through a very arduous process of taking up the general point of view, and also that it is unclear how much correction or regulation is necessary for us to be able to make moral judgments. If such an inaccessible point of view were essential for moral sentiment to arise, then most of the moral judgments we make in daily life would be not moral judgments at all—not even “unjust” moral judgments. In other words, the extension of moral judgment would become too narrow.

Garrett's argument is persuasive, and it is unlikely that all individual moral judgments or moral sentiments are only caused by taking up the general point of view.⁹ Therefore, it would have to be said that the etiological attempt to specify moral sentiments by invoking the general point of view as their necessary cause has failed.¹⁰ On the other hand, the phenomenological identification of moral sentiment, which Garrett attributes to Hume, is also not plausible as a matter of fact, and, as we saw in the previous subsection, Hume himself admits that moral and interested sentiments are so phenomenologically similar that they are routinely confused.

From the above discussion, it can be concluded that all the basic attempts to solve the Moral Attitude Problem in Hume's moral theory have been unsuccessful. Thus, we are brought back to our initial question. What exactly is the “moral sentiment” that Hume refers to? Is there really any characteristic common and unique to all moral sentiments that allows them to adequately determine the extension of moral judgment?

This problem, which already seems difficult to solve, becomes even more puzzling when we realize the nature of moral judgments based on the “general point of view.” Hume presumably regards judgments of that sort as exemplary moral judgments,¹¹ but remarkably,

⁸ Abramson (1999) describes in great detail how complicated and endless these corrections and regulations can be.

⁹ This point is discussed again in Section 5.

¹⁰ There is another (more permissive) etiological answer that at least “sympathy” (but not the general point of view) is necessary as a cause for the generation of moral sentiments (Carlson 2014: 82). However, it is also ambiguous even whether Hume thought that sympathy was necessary for all moral judgments (T 3.3.1.27–29; cf. Garrett 2001b: 213; Carlson 2014: 93). Also, as we will discuss below, not even sentiment may be necessary for moral judgments in the first place, and if that were the case, sympathy, the principle that transmits sentiment, would not be necessary either.

¹¹ Hume says that “they [= interests and pleasures perceived from the general point of view] [...] are alone admitted in speculation as the standard of virtue and morality. They alone produce that particular feeling or sentiment, on which moral distinctions depend” (T 3.3.1.30).

such judgments are often described as having little sentimental aspect. If this means that sentiment is not the indispensable component of moral judgments, then it follows that there is no such thing as the moral sentiment we seek—the kind of sentiment that constitutes every individual moral judgment and thus determines the extension of moral judgment. However, in my view, it is this seemingly perplexing possibility that implies the very nature of what Hume calls “moral sentiment.” In the next section, therefore, I will examine moral judgments based on the general point of view in more detail.

3. Moral Judgment from the General Point of View: Dilution of Sentimental Aspect and Emphasis on the Role of Language

It has often been pointed out and problematized by Hume scholars that moral judgments based on the general point of view may lack sentimental aspects (cf. Cohon 2008: 127). If there are non-sentimental but genuine moral judgments, then Hume’s moral sentimentalism, which holds that moral judgments are constituted by sentiments, would be inconsistent. Put another way, there would be no such thing as a moral sentiment that could determine the extension of moral judgment and thus be the answer to the Moral Attitude Problem. In this section, I will confirm that the introduction of the “general point of view” leads to this *dilution of the sentimental aspect* of moral judgments and instead to an emphasis on the role of *language* in moral judgments, and then I will show that this implies a certain *fictionality* of “moral sentiment” in Hume’s moral theory.

First, as seen in the previous subsection, for one to take up the general point of view, one sometimes needs to sympathize with those who are remote from oneself on the same level as with one’s neighbors (for example, when one is in England and the object of one’s judgment is in China), but it is unclear how one can do so because sympathy was originally introduced as a mechanism that works more strongly and easily with those who are in some sense closer to us. In fact, Hume says the following about some moral judgments based on the general point of view:

Our servant, if diligent and faithful, may excite stronger sentiments of love and kindness than *Marcus Brutus*, as represented in history; but we say not upon that account, that the former character is more laudable than the latter. We know, that were we to approach equally near to that renown’d patriot, he wou’d command a much higher degree of affection and admiration. (T 3.3.1.16)

We blame equally a bad action, which we read of in history, with one perform'd in our neighbourhood t'other day: The meaning of which is, that we know from reflection, that the former action wou'd excite as strong sentiments of disapprobation as the latter, were it plac'd in the same position. (T 3.3.1.18)

The “affection and admiration” and “sentiments of disapprobation” mentioned here do not seem to be something that is *felt* by sympathy, but merely something that is comprehended through imagination. If to judge based on such sentiments is to judge from the general point of view, such judgments would *not* be based on the observer's own *sentiments*.

In addition to this dilution of the sentimental aspect of moral judgments, the introduction of the general point of view also brings an emphasis on the role of *language*. As in the first quote, after the introduction of the general point of view, expressions like “we *feel* this way but do *not say* so” begin to appear (T 3.3.1.15; T 3.3.1.21). And the following passage is decisive.

[T]hese variations [of sentiments] we regard not in our general decisions, but still apply the terms expressive of our liking or dislike, in the same manner, as if we remain'd in one point of view. Experience soon teaches us this method of correcting our sentiments, or at least, of correcting our language, where the sentiments are more stubborn and inalterable. (T 3.3.1.16)

So, there are two possible outcomes of adopting the general point of view. One is that the sentiment—how to feel—itself is corrected according to the general point of view; and the other is that the sentiment is not corrected and only the language, or the application of moral terms—what to say—is corrected according to the general point of view. The important point is that if the product of either correction can be regarded as moral judgment (as Hume seems to think), then sentiments play no fundamental role in the determination of the extension of moral judgment. This remains so whether the general point of view is essential to moral judgment or not. In any case, if both sentiment-based and non-sentiment-based judgments can be moral judgments, then this is the case. If this is true, then the Moral Attitude Problem for Hume has entered a whole new phase.

However, perhaps aware of this point, Hume suggests that moral judgments based on the general point of view still have a sentimental underpinning.

'Tis seldom men heartily love what lies at a distance from them, and what no way redounds to their particular benefit; as 'tis no less rare to meet with persons, who can pardon another any opposition he makes to their interest, however justifiable that

opposition may be by the general rules of morality. Here we are contented with saying, that reason requires such an impartial conduct, but that 'tis seldom we can bring ourselves to it, and that our passions do not readily follow the determination of our judgment. This language will be easily understood, if we consider what we formerly said concerning that *reason*, which is able to oppose our passion; and which we have found to be nothing but a general calm determination of the passions, founded on some distant view or reflection. [...] This [sympathy with people designated by the general point of view] is far from being as lively as when our own interest is concern'd, or that of our particular friends; nor has it such an influence on our love and hatred: But being equally conformable to our calm and general principles, 'tis said to have an equal authority over our reason, and to command our judgment and opinion. (T 3.3.1.18)

That is to say, the pleasure or pain perceived from the general point of view, however faint, is sufficient to activate the “calm determination of passions” or “calm and general principles” (we will see shortly why these are called “reason”). And this is enough to control the use of language.

And tho' the *heart* does not always take part with those general notions, or regulate its love and hatred by them, yet are they sufficient for discourse, and serve all our purposes in company, in the pulpit, on the theatre, and in the schools. (T 3.3.3.2)

In other words, the correction of language, without the correction of sentiments, through the general point of view is still supported by some kind of faint sentiment.

Given Hume's definitions of the “general point of view” and each person's “peculiar point of view” as seen in Section 2.3, the situation is probably as follows: When X considers Y's character from the general point of view, there emerge in X's mind, by virtue of the sympathy that constitutes the general point of view, the faint pleasures and pains that Y's character brings to Y and those around Y. However, there also still remain in X's mind, due to X's peculiar point of view, the lively pleasures and pains that Y's character brings to X and those with whom X naturally sympathizes. Then, if X somehow comes to entertain more lively feelings or sentiments from the former faint pleasures and pains than from the latter, that is the correction of sentiments. And even if not, X can still entertain faint sentiments from the former pleasures and pains and apply moral terms (e.g., “virtuous” or “vicious”) based on those sentiments, which is the correction of language.¹²

¹² This interpretation is roughly the same as that of Cohon (2008: 139–141), but unlike her, I doubt that these “faint sentiments” are genuine sentiments, as discussed below. Also, note that one flaw in this interpretation is that it does not explain the “somehow,” that is, how we can feel lively toward

However, these faint “feelings” or “sentiments” do not seem to be what we would normally understand as sentiments. Hume says the following of “calm passions” (T 2.3.3.9) in his theory of passions:

’Tis natural for one, that does not examine objects with a strict philosophic eye, to imagine, that those actions of the mind are entirely the same, which produce not a different sensation, and are not immediately distinguishable to the feeling and perception. [...] 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. [...] 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. (T 2.3.3.8)

In short, calm passions are phenomenologically so weak that (1) they are known by their effects (speech or behavior, typically) rather than by their direct feeling, and (2) they are easily confused with the judgments of reason, which is the capacity to judge truth or falsehood. These two characteristics would also apply to some moral sentiments arising from the general point of view. This is because, in the earlier quotation, Hume says, clearly based on this passage, that “reason,” or in fact “a general calm determination of the passions,” requires us to make moral judgments from the general point of view.

Now, the implication that some moral sentiments have these two characteristics is significant. Let us focus here on the first characteristic.¹³ The faint moral sentiments at issue are not specified by their phenomenological feelings (which Hume initially relied on), but rather by their perceptible effects—what we say or do. Given that, for Hume, to have moral sentiments is to make moral judgments, the most obvious effect of them would be the speech act of applying moral terms to someone in some way (e.g., uttering, “He is vicious”). In this light, we can see why Hume insists that even when only the language is corrected, some faint moral sentiment is still entertained. Hume retrospectively assumes that the corresponding moral sentiment existed from the result that the application of moral terms has been corrected or controlled according to the general point of view. This means that “moral sentiments” for Hume, even if they are the essence of moral judgment and thus determine its extension, are at

the faint pleasures and pains. King argues that the main problem with the *Treatise* is precisely that it fails to causally explain this “correction of the sentiments” (King 1976: 350–353).

¹³ The second characteristic will be touched upon briefly at the end of this paper.

least sometimes not perceived in themselves, but are merely retrospectively hypothesized to exist from a certain kind of speech act. If so, the extension of moral judgment is determined, at least in part, by language, not by sentiment. This is a remarkable reversal.

However, there is room to argue against this interpretation as follows. Although those who make moral judgments using moral terms from the general point of view often do not entertain real moral sentiments, the very moral terms used there instead express lively moral sentiments by their linguistic function.¹⁴ In other words, even when only language is corrected according to the general point of view, the language itself shoulders the sentimental aspect of the moral judgment. “[L]anguage [...] must invent a peculiar set of terms, in order to express those universal sentiments of censure or approbation” (EPM 9.8). Paradigmatic moral judgments are judgments constituted by lively, perceptible, and real moral sentiments, and only insofar as expressing such moral sentiments are verbal judgments with no conscious sentiments also regarded as moral judgments. If this is the case, then what fundamentally determines the extension of moral judgment is still sentiment, not language.

In response to this objection, I argue that moral language, or the set of moral terms (which Hume calls “another language”¹⁵), cannot possibly have the linguistic function of expressing lively sentiments, given the process by which it is established. Rather, moral language is intrinsically tied to faint “sentiments” stripped of liveliness. To show this, in the next section, I will focus on the genealogical aspect of Hume’s theory of moral judgment and trace in detail how moral language is created. This exploration will eventually reveal not only what role is entrusted to moral language, but also why the “moral sentiment” that constitutes moral judgment and determines its extension must be fictional.

¹⁴ The mental state that constitutes moral judgment is logically distinct from the mental state that moral terms express as their function (Joyce 2009: 38–39). For example, even if one does not feel conscious displeasure with someone’s wrong action, one can mechanically say, “What he did was wrong” to *express one’s displeasure*, and this will often succeed in conveying a nonexistent displeasure to the hearers.

¹⁵ Hume says, “When a man denominates another his *enemy*, his *rival*, his *antagonist*, his *adversary*, he is understood to speak the language of self-love, and to express sentiments, peculiar to himself, and arising from his particular circumstances and situation. But when he bestows on any man the epithets of *vicious* or *odious* or *depraved*, he then speaks another language [...]” (EPM 9.6).

4. Hume's Genealogy of Moral Judgment: The Establishment of the General Point of View and the Creation of Moral Language

In Hume's moral theory, the creation of moral language is always mentioned in the context of explaining the general point of view. In his *Treatise*, however, it is only hinted at.

[It is] impossible we cou'd ever converse together on any reasonable terms, were each of us to consider characters and persons, only as they appear from his peculiar point of view. (T 3.3.1.15)¹⁶

[It were] impossible we cou'd ever make use of language, or communicate our sentiments to one another, did we not correct the momentary appearances of things, and overlook our present situation. (T 3.3.1.16)

So, at least the use of moral language and the communication through it are causally dependent on everyone leaving one's own peculiar point of view and taking up the general point of view, common to everyone. And in his *Enquiry*, Hume explicitly refers to the creation of moral language in the same context.

General language, therefore, being formed for general use, must be moulded on some more general views, and must affix the epithets of praise or blame, in conformity to sentiments, which arise from the general interests of the community. (EPM 5.42)¹⁷

Hence, there is no doubt that moral language is created based on the general point of view.

Then, Hume's "general point of view" has two radically different roles related to moral language. One is to control the use of existing moral language in making moral judgments in everyday life, as discussed in the previous section, and the other is to create moral language itself in the more primitive circumstances of humanity. In fact, it is crucial to distinguish between these two roles to plausibly interpret Hume's descriptions of the *process by which we arrive at* the general point of view as follows.

Our situation, with regard both to persons and things, is in continual fluctuation; and a man, that lies at a distance from us, may, in a little time, become a familiar acquaintance. Besides, every particular man has a peculiar position with regard to others; and 'tis impossible we cou'd ever converse together on any reasonable terms, were each of us

¹⁶ See also T 3.3.3.2.

¹⁷ See also EPM 9.6–8.

to consider characters and persons, only as they appear from his peculiar point of view. In order, therefore, to prevent those continual contradictions, and arrive at a more stable judgment of things, we fix on some steady and general points of view; and always, in our thoughts, place ourselves in them, whatever may be our present situation. (T 3.3.1.15)

When we form our judgments of persons, merely from the tendency of their characters to our own benefit, or to that of our friends, we find so many contradictions to our sentiments in society and conversation, and such an uncertainty from the incessant changes of our situation, that we seek some other standard of merit and demerit, which may not admit of so great variation. Being thus loosen'd from our first station, we cannot afterwards fix ourselves so commodiously by any means as by a sympathy with those, who have any commerce with the person we consider. (T 3.3.1.18)

The intercourse of sentiments, therefore, in society and conversation, makes us form some general inalterable standard, by which we may approve or disapprove of characters and manners. (T 3.3.3.2)

To summarize: Insofar as each observer shows a positive or negative reaction to a person based on the pleasures and pains that the person brings to each observer and those whom each observer naturally sympathizes with, these pleasures and pains are relative to each observer's position at any given time so that their reactions to the same person will perpetually disagree in their social interaction and conversation. Observers who have experienced such a predicament for some time, trying to avoid such disagreement, begin to leave their own peculiar point of view, to sympathize with the pleasures and pains that the person in question brings to herself and those around her (that is, to take up the general point of view), and to show the same reaction based on these common pleasures and pains to each other in society and conversation. Such common pleasures and pains (perceived from the general point of view) then become the general criterion for evaluating people.

First of all, I would like to point out that if this were a description of the way people who can already use moral language make moral judgments in their daily lives, it would not be plausible at all. In everyday life, we often agree in our moral judgments about a person without the effort described here, and even when we disagree, we might begin to dispute, but we can still understand what each other is saying and continue the conversation (which is why we can dispute). If the disagreement is intolerably uncomfortable, one can simply end the conversation or leave the room. Also, as Hume himself says, when we read the biographies of foreign

historical figures, we often praise or blame them for the happiness or misfortune they have brought to their country (cf. T 3.3.1.16; T 3.3.1.18). In this case, however, we need not go through any social interaction and conversation between reading the biography and evaluating the person, and moreover, the evaluation is already based on the “general point of view.” Furthermore, it would be strange if “some general inalterable standard” were sought anew with each attempt to make a moral judgment.

Given all this, Hume’s descriptions of the process leading to the general point of view through “contradictions” in “society and conversation” are far more plausible when interpreted not as concerning our moral judgment practices in everyday life, but as concerning the behavior of people in a much more primitive situation before the creation of moral language. In fact, as we have already seen, this process is tied to the creation of moral language through its final product, “some more general views.” It would be even stranger if moral language were created every time we tried to make moral judgments in our daily lives.

So, this is the process in which the general point of view has been established in the long history of humanity, as something so constant that it is linked to the creation of moral language. But how exactly is it established? Elucidating this point will clarify what kind of sentiment is connected with moral language.¹⁸

According to Hume, language is created through human “convention,” which is nothing but “a general sense of common interest” (T 3.2.2.10; cf. EPM app. 3.7–8). Therefore, for moral language to be created, it is necessary that people have a “common interest” in its creation and that they are all aware of it.¹⁹ This means that people have a common interest in establishing and fixing the general point of view in the form of moral language and that they must be aware of this interest. So, what exactly is this common interest, and how do people become aware of it?

As to the first question, Hume gives a direct answer in the above quotes. Namely, the common interest in establishing the general point of view is “to prevent those continual contradictions [to our sentiments in society and conversation], and to arrive at a more stable judgment of things.” But what does this “contradiction” mean? It motivates people to leave

¹⁸ In his *Enquiry*, Hume introduces the concept of “sentiment of humanity” and simply links this to moral language (EPM 9.8). Still, since Hume does not provide a clear definition of this concept, I do not think that invoking it will help to clarify the nature of moral language and sentiment.

¹⁹ In the course of explaining the origin of justice, which is also established through human convention in order to form and sustain society, Hume says, “But in order to form society, ’tis requisite not only that it be advantageous, but also that men be sensible of its advantages” (T 3.2.2.4). I obtained the idea of basing the general point of view on convention from Mackie (1980: 123).

their initial and comfortable viewpoints of self-interest and natural sympathy. Therefore, the contradiction must be distasteful to people.

On this point, Korsgaard's (1999: 23–25) interpretation seems persuasive. For the sake of brevity, let us assume that two observers show each other their reactions to a person in their conversation. If one shows a positive reaction and the other a negative reaction, the two observers would sympathize with each other, resulting in the coexistence of conflicting sentiments toward the same person in each observer's mind. Hume describes this kind of conflict as unpleasant (T 2.1.11.19), and so each observer has the motivation to avoid this unpleasantness. This is certainly a plausible interpretation. Notably, if, as I believe, the scene in question is one before the creation of moral language, the "contradictions" cannot be a linguistic conflict or dispute, so it is expedient to appeal to a conflict of sentiments that are somehow intercommunicated.

In addition, we cannot simply abandon "society and conversation" even if they frequently cause intolerable discomfort because:

[C]ompany is naturally so rejoicing, as presenting the liveliest of all objects, *viz.* a rational and thinking Being like ourselves, who communicates to us all the actions of his mind; (T 2.2.4.4)

A perfect solitude is, perhaps, the greatest punishment we can suffer. Every pleasure languishes when enjoy'd apart from company, and every pain becomes more cruel and intolerable. (T 2.2.5.15)

To avoid social isolation, people must continue to interact with others to some degree and, therefore, are forced to strive for consensus in the evaluation of persons so that conflict and discomfort do not occur in such interaction. (In contemporary large societies, it may be possible for people to associate only with those who share similar beliefs and values, but this would be impossible in small communities, especially primitive ones that existed before the creation of moral language.)

However, this discomfort of the "contradiction" alone does not yet motivate people to try to establish the general point of view itself. At this point, people are still seeking a means to achieve the goal of consensus and eliminate the discomfort. Only after realizing that the means is to take up the general point of view will they be motivated to establish it (cf. T 2.3.3.3).

Unfortunately, Hume does not directly explain how this realization arises. Thus, it has been a traditional interpretive question as to why people move toward the general point of view

(e.g., Korsgaard 1999; Cohon 2008: Ch. 5). Often noted in this context is the following passage.

When experience has once given us a competent knowledge of human affairs, and has taught us the proportion they bear to human passion, we perceive, that the generosity of men is very limited, and that it seldom extends beyond their friends and family, or, at most, beyond their native country. Being thus acquainted with the nature of man, we expect not any impossibilities from him; but confine our view to that narrow circle, in which any person moves, in order to form a judgment of his moral character. When the natural tendency of his passions leads him to be serviceable and useful within his sphere, we approve of his character, and love his person, by a sympathy with the sentiments of those, who have a more particular connexion with him. [...] And tho' this advantage or harm be often very remote from ourselves, yet sometimes 'tis very near us, and interests us strongly by sympathy. This concern we readily extend to other cases, that are resembling; (T 3.3.3.2)²⁰

Since most people only care about the pleasures and pains of those around them, or their “narrow circle,” the observer, having learned this from experience, no longer expects any person to care about the pleasures and pains of those who are remote from the person, and as a result, when considering the character of a person, the observer’s gaze naturally turns to the person’s narrow circle. Furthermore, the narrow circle of the person in question is sometimes located close to the observer, in which case the observer naturally sympathizes with that narrow circle and receives lively pleasures and pains, and thus can react positively or negatively to the person based on those lively feelings. Then, those who have experienced this will come to apply the same type of reaction based on the narrow circle to cases where someone’s narrow circle is far away from them and thus its pleasures and pains appear only faint to them.²¹ By the way, in the “narrow circle” of a person, there is the person herself and the people around her, so to sympathize with their pleasures and pains is nothing but to take up the general point of view.

However, this explanation (even if we add the various reasons often proposed by Hume scholars as reinforcing this) seems still insufficient for people establishing the general point of

²⁰ See also T 3.2.1.18.

²¹ This sequence of the process is often interpreted as involving the function of “general rules” (T 3.3.1.20) that make people imagine similar effects from similar causes even if the causes differ in important respects (cf. Magri 1996: 236–239).

view as so constant and privileged that it becomes the cornerstone of a new language.²² This is because the circumstances described here do not seem to provide a motivation powerful enough to make us “forget” (T 3.3.1.30; T 3.3.3.2) the intense pleasures and pains based on our self-interest and natural sympathy. For example, if a person’s narrow circle includes our friends or enemies, how can we be prevented from valuing that person higher or lower than otherwise? Or, when one person’s narrow circle is close to us and another person’s narrow circle is far away, how can we be prevented from valuing the former higher or lower than the latter? Certainly, to “expect not any impossibilities from him” may be a reason for us to look to that person’s narrow circle, but it is not a reason for us to forget our own interests and natural feelings. Indeed, even after moral language has been created, it is still difficult for us to do so, that is, to apply moral terms properly to our “enemies” (T 3.1.2.4), as we saw in Section 2.2.

In short, this type of explanation, which appeals to the habit or rationality of human nature, can show why people tend to look to the pleasures and pains of the narrow circle in question, but it cannot show why people focus *exclusively* on those pleasures and pains in a particular context. The only plausible explanation for the latter, I believe, is still that people have somehow realized that unless they do so, they cannot prevent the “contradiction.” In fact, Hume himself always refers to this point when he tries to explain why people can forget about pleasures and pains other than those of the narrow circle.

One may, perhaps, be surpriz’d, that amidst all these interests and pleasures, we shou’d forget our own, which touch us so nearly on every other occasion. But we shall easily satisfy ourselves on this head, when we consider, that every particular person’s pleasure and interest being different, ’tis impossible men cou’d ever agree in their sentiments and judgments, unless they chose some common point of view, from which they might survey their object, and which might cause it to appear the same to all of them. (T 3.3.1.30)

We are quickly oblig’d to forget our own interest in our judgments of this kind, by reason of the perpetual contradictions, we meet with in society and conversation, from persons that are not plac’d in the same situation, and have not the same interest with ourselves. The only point of view, in which our sentiments concur with those of others,

22 For example, Korsgaard (1999: 26–35) explains this point by appealing to the nature of love/hatred that accompanies the evaluation of a person, but in Hume, love/hatred itself has no motivational power, and moreover, benevolence/anger that accompanies love/hatred is the desire for the happiness/misery of the loved/hated one (T 2.2.6.3), which would only further bias the evaluative sentiment. As I will discuss later, I believe that the general point of view is established where all of these passions and sentiments are ignored.

is, when we consider the tendency of any passion to the advantage or harm of those, who have any immediate connexion or intercourse with the person possess'd of it. (T 3.3.3.2)

After all, the only motivation powerful enough to make us forget our self-interests and natural sympathies is the common desire to avoid this “contradiction.” Therefore, unless this desire is directly connected to taking up the general point of view, we will never voluntarily and collaboratively pursue the general point of view when evaluating people and thus establish it as a permanent and common criterion that leads to the creation of moral language.

What, then, links this desire to the general point of view? I think it can be nothing more than the repeated experience in the long history of mankind that *everyone in an exchange happens to take up the general point of view and, as a result, an agreement of reactions and an avoidance of contradictions are actually realized*. Only when this happens do people learn that taking up the general point of view is the most reliable means of satisfying their strong desire to avoid contradictions, and they come to desire this point of view directly, even if it goes against their self-interests and natural sympathies.²³

It should be noted, however, that an occurrence of such a coincidental “agreement” of people’s reactions to a certain person does not immediately guarantee that the agreement was based on the general point of view. For example, in a very small community, where all the residents know each other, their reactions to the same person may sometimes agree simply by letting their self-interest and natural sympathy take over. But this kind of agreement does not last. As the community expands, as more and more strangers come into the mix, as more and more observers from different perspectives interact, it is no longer possible to achieve agreement by each observer’s self-interest and natural sympathy. To achieve agreement here, we must react to the person in question based solely on the pleasures and pains of that person and those around her, even when those people are far from us or are our enemies. That is to say, we must take up the general point of view. As Hume said, “The more we converse with mankind, and the greater social intercourse we maintain, the more shall we be familiarized to these general preferences and distinctions” (EPM 5.42). The reason for this, in my interpretation, is that in a larger community, the coincidental agreement necessary to avoid the “contradictions” can only occur reliably by taking up the general point of view.

²³ Similarly, in his theory of justice, Hume states that it is through the *actual experience of society*, a family community consisting of both sexes and children, that people realize the advantages of society (T 3.2.2.4).

What is important, anyway, is that this kind of agreement actually occurs throughout the history of humankind. Only then are people motivated to establish the general point of view and, thus, to create moral language. But when, exactly, can this kind of agreement occur? I ask this because merely everyone focusing on the “narrow circle” of the person in question seems to be not enough to achieve actual agreement. What is additionally needed, I think, is (1) sharing the perception of pleasures and pains and (2) ignoring the liveliness of pleasures and pains.

(1) Sharing the perception of pleasures and pains: If the observers did not have a shared perception of the pleasures and pains that the person in question was bringing to herself and those around her, their reactions to the person based on those pleasures and pains could never be the same. In fact, it is ambiguous what kind of people “those around her” or her “narrow circle” includes.²⁴ But in any case, if, by chance, the observers happen to focus on the same people and perceive the same pleasures and pains, then at least the basis for agreement is secured.

(2) Ignoring the liveliness of pleasures and pains: This is far more important for our purpose. It relies on the following assertion by Hume.

[T]ho’ sympathy be much fainter than our concern for ourselves, and a sympathy with persons remote from us much fainter than that with persons near and contiguous; yet we neglect all these differences in our calm judgments concerning the characters of men.
(T 3.3.3.2)

Namely, the various pleasures and pains that the person in question brings to various people are perceived by each observer with varying degrees of liveliness, depending on the distance between each observer and these people. However, when we take up the general point of view, we ignore all such differences in liveliness. I think this must mean that we *ignore* not only the *differences* in liveliness but also the *liveliness itself*. This is because, for example, if the group of pleasures that observers A and B perceive in common is as a whole near (in some sense) to A and far (in some sense) from B, then if A and B merely ignore the differences in the liveliness of their respective pleasures, A will perceive homogeneous but *lively* pleasures while B will perceive homogeneous but *faint* pleasures. Of course, such pleasures differ in their liveliness

²⁴ There is a very big difference between “friends and family” and “native country,” which Hume mentions as the “narrow circle,” and it is also strange, at least in our time, that we cannot expect humans to care at all for people from other countries (T 3.3.3.2). Moreover, this range can even be variable, depending on various conditions, such as the type of virtue, the influence of the person in question, the social context, etc. (cf. Baier 2006; Brown 1994: 34n.12; Abramson 1999: 344).

between A and B, so they are still not “pleasure, which appears the same to every spectator” (T 3.3.1.30). Moreover, according to Hume, such differences in the liveliness of perceived pleasure or pain naturally lead to differences in the *kind* of reaction of the perceiver: (a) when the liveliness of the perceived pleasure is strong enough, the perceiver feels pleasure himself through sympathy; (b) when the degree of liveliness is not so strong, the perceiver feels pain by comparison with his own state; and (c) when the liveliness is significantly weak, the perceiver feels nothing himself (T 3.3.2.4–5; cf. Abramson 1999: 339–341). Hence, for the observers’ reactions to agree, each observer must ignore the liveliness itself of his or her perceived pleasure or pain and then somehow react based on that pleasure or pain.

Now, if by chance all of the above are realized, surely the reactions of all observers will agree. Since all observers are looking at the same people around the person in question, perceiving the same kinds of pleasures and pains, and ignoring their liveliness, their reactions based on such pleasures and pains must be identical. And even if such agreement is coincidental at first, as a method of achieving it is learned and repeated, the method or the general point of view will finally be established. Based on this point of view, moral language is created.

At last, we can now answer the question posed in the previous section. The moral sentiment connected with moral language must be the kind of sentiment based on *pleasure or pain whose liveliness has been ignored* since it is only such sentiments that have realized agreements of reactions and led to the establishment of the general point of view and moral language. Therefore, what moral language expresses cannot be lively sentiments, but rather “sentiments” that have been stripped of their emotional aspects, have no phenomenologically distinctive feeling, and can only be retrospectively hypothesized to exist from the very use of moral language. When only our language, not our sentiments, is corrected according to the general point of view, we are making moral judgments that have nothing to do with substantive and real sentiments. Hume acknowledges the existence of such moral judgments.

Let me answer one supposed objection here. Certainly, Hume says the following.

[H]ad not men a natural sentiment of approbation and blame, it cou’d never be excited by politicians; nor wou’d the words *laudable* and *praise-worthy*, *blameable* and *odious*, be any more intelligible, than if they were a language perfectly unknown to us, (T 3.3.1.11)²⁵

But I need not deny this. Indeed, moral language has its origins in human sentiments. If humankind had no natural sentiments, positive or negative, toward others, the “contradictions”

²⁵ See also T 3.2.2.25.

in society and conversation could not have occurred, and thus neither the general point of view nor moral language would have been relevant to us. However, this does not mean that those natural sentiments are directly connected to moral language. In fact, as we have seen, Hume ultimately separates them. For otherwise he could not have fully explained the nature of moral judgments. Moral judgment has its origin in sentiment, but it is completed when that sentiment is disregarded and that disregard is entrusted to language. Hume's emphasis on the connection between moral judgments and sentiments is consistent with my interpretation, as long as we regard his emphasis not as a psychological analysis of individual moral judgments, but as a genealogical analysis of moral judgments in general.

5. The Primacy of Language over Sentiments in Hume's Moral Philosophy

The above discussion has shown that, in Hume's moral theory, at least some moral judgments are not constituted by perceptible moral sentiments nor express them, but are merely hypothesized to be underpinned by "moral sentiments" retrospectively from the use of moral language. Therefore, the extension of moral judgment is determined, at least in part, by language, not sentiments.

But I would like to offer here an interpretation that goes one step further. It is that, in Hume's moral theory, the extension of moral judgments is *entirely* determined by moral language, not by moral sentiments, in the first place. This may seem a surprising assertion, but Hume himself suggests such an idea in his *Enquiry*.

The distinction, therefore, between these species of sentiment being so great and evident, language must soon be moulded upon it, and must invent a peculiar set of terms, in order to express those universal sentiments of censure or approbation[...]. VIRTUE and VICE become then known: Morals are recognized: Certain general ideas are framed of human conduct and behaviour: Such measures are expected from men, in such situations: This action is determined to be conformable to our abstract rule; that other, contrary. (EPM 9.8)

The "universal sentiments" correspond to the sentiments arising from the general point of view, based on which moral language is created so that morality, including the distinction between virtue and vice, comes to be recognized. And indeed, Hume uses such language as a reliable guidepost for his moral theory.

The very nature of language guides us almost infallibly in forming a judgment of this nature; and as every tongue possesses one set of words which are taken in a good sense, and another in the opposite, the least acquaintance with the idiom suffices, without any reasoning, to direct us in collecting and arranging the estimable or blameable qualities of men. (EPM 1.10)²⁶

This circumstance would also apply to the moral theory in his *Treatise*, even if Hume himself is unaware of it. Insofar as Hume (and probably all moral philosophers) is forced to consider moral judgments within the scope of what can be indicated by moral terms such as “virtue” and “vice,” the extension of moral judgment is determined in advance by moral language (cf. King 1976: 356). And since, as we have seen in this paper, there are no perceptible sentiments that underpin all and only such judgments, it is moral language that essentially determines the extension of moral judgment. Hume was content to argue that even when sentiments cannot be corrected, language can be corrected, precisely because the latter—how and which moral terms are applied—is crucial to moral judgments, and the former—what kind of sentiments are entertained—is merely incidental to them.

It should be noted here that the fact that someone has made a moral judgment using moral terms, which were created based on the general point of view, does *not guarantee* that the person has *actually* taken up the general point of view, that is, focused exclusively on the pleasures and pains of the narrow circle in question and ignored their liveliness. Sometimes people may unconsciously confuse their own or their relatives’ pleasures and pains with those of the narrow circle, and sometimes they may consciously abuse moral terms to defame their opponent.

[T]hough much of our friendship and enmity be still regulated by private considerations of benefit and harm, we pay, at least, this homage to general rules, which we are accustomed to respect, that we commonly pervert our adversary’s conduct, by imputing malice or injustice to him, in order to give vent to those passions, which arise from self-love and private interest. (EPM 9.8 n.57)

Such confusion and abuse are possible because, regardless of whether we are actually taking up the general point of view or not, the moral terms we apply themselves express, by their linguistic function, *something perceived from the general point of view*.²⁷ It is in the gap

²⁶ See also Hume ([1777] 1985: 229).

²⁷ For the moment, I believe this “something” is either a “universal but fictional sentiment” or an “objective moral property.”

between this function of moral language and our actual attitudes that confusion and abuse can exist.²⁸ If only our actual attitudes—whether we take up the general point of view, whether we entertain a particular kind of sentiment—mattered, we would either be making moral judgments or we would not, and biased, unjust, or insincere moral judgments would not exist in the first place.

This distinction between the linguistic function and our attitudes provides another answer to the Garrett–Brown controversy over the status of the general point of view, which we saw in Section 2.3. I think that to make moral judgments (or to be assumed to have “moral sentiments”), we need not actually take up the general point of view, but we do need to use moral terms with an understanding of their function of expressing something perceived from the general point of view.²⁹ Actually taking up the general point of view is neither a distinctive cause nor a causal precondition common to all individual moral judgments (thus, the problems Garrett pointed out do not arise). All that is required to make moral judgments is to be competent to use moral terms, that is, to understand what it is like to view an object from the general point of view and to connect that understanding with moral terms.

Finally, let us see what happens to the Moral Attitude Problem for Hume in light of the interpretation of this paper. In conclusion, there is no moral attitude that determines the extension of moral judgment without excess or deficiency. This is because, as we have seen, in Hume’s moral theory, it is language, not sentiment, that essentially determines the extension of moral judgment. When we use moral language in a particular way, we are making moral judgments, nothing more and nothing less. Although Hume inserts the underpinning of “moral sentiments” there, such sentiments themselves do not play any essential role in determining the extension of moral judgment, since the sentiments can only be retrospectively assumed from the use of moral language. Certainly, some of our moral judgments may involve intense and perceptible sentiments (such as when the person in question is close to us), but that is only incidental to the fact that they are moral judgments.

²⁸ As an analogy, you can make an improper or false promise by saying, “I promise to do X,” when you mistakenly believe you can do X while you actually cannot, or when you have no intention of doing X in the first place.

²⁹ If a person who does not know the meaning of the word “promise” were to utter “I promise,” he would not have made a promise.

6. Conclusion

In this paper, I have shown that Hume, who initially emphasized the connection between moral judgment and sentiment, eventually made that sentiment diluted and fictional, and then implied instead the primacy or logical precedence of language in moral judgment. This shift could be taken to indicate that Hume failed in his psychological analysis of moral judgments (as King says), but I would rather see it as presenting a fascinating genealogical analysis that reveals that in human history, moral judgment was established when we disregarded our sentiments and entrusted that disregard to moral language. King states that Hume in the *Enquiry* relies “on institutions and practices to shed light on a freshly conceived *cultural*, rather than natural, Science of Man” (King 1976: 357), and in my view, the third book of his *Treatise* depicts in detail how those “institutions and practices” have been established from human nature. By tracing this process, we can go beyond merely accepting them as given and understand why they had to be so, namely, why real and perceptible sentiments had to be excluded from moral language and thus from moral judgments.

Now, if the interpretation of this paper is correct, many further issues arise that need to be discussed. First, if moral judgments are not constituted by substantive sentiments, what happens to Hume’s emphasis on the practicality of moral judgments, that is, the characteristic that moral judgments motivate the judge? (T 3.1.1.15) Second, in the quotation introduced in Section 3, Hume said that ordinary people confuse “calm passions” with “determinations of reason.” This confusion would also apply to “moral sentiments.” Are then moral judgments that are supposedly underpinned by such sentiments conceptually a belief or representational state, and truth-apt? If so, what are their truth conditions? Can they ever actually be true? Third, since the general point of view was established through the experience of the agreement of people’s reactions, it seems that moral judgments from that point of view are conceptually assumed to agree among people, but can there be universal agreement in moral judgments? Fourth, is the general point of view, established simply to avoid conversational discomfort, really the point of view we ought to adopt in the first place? I would like to address at least these issues in the future.

Acknowledgments

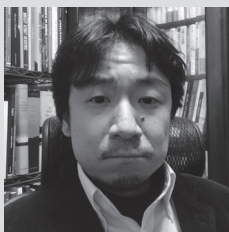
I would like to thank the anonymous reviewers for their helpful comments. This work was supported by JSPS KAKENHI Grant Number JP21K19942.

References

- Abramson, K. (1999). Correcting Our Sentiments about Hume's Moral Point of View. *The Southern Journal of Philosophy*, 37, 333–361.
- Baier, A. C. (2006). How Wide is Hume's Circle? (A question raised by the exchange between Erin I. Kelly and Louis E. Loeb, *Hume Studies*, November 2004). *Hume Studies*, 32(1), 113–117.
- Björnsson, G., & McPherson, T. (2014). Moral Attitudes for Non-Cognitivists: Solving the Specification Problem. *Mind*, 123(489), 1–38.
- Blackburn, S. (1993). *Essays in Quasi-Realism*. Oxford University Press.
- Brown, C. (1994). From Spectator to Agent: Hume's Theory of Obligation. *Hume Studies*, 20(1), 19–36.
- Brown, C. (2001). Is the General Point of View the Moral Point of View? *Philosophy and Phenomenological Research*, 62(1), 197–203.
- Carlson, A. (2014). The Moral Sentiments in Hume's Treatise: A Classificatory Problem. *Hume Studies*, 40(1), 73–94.
- Cohon, R. (2008). *Hume's Morality: Feeling and Fabrication*. Oxford University Press.
- Garrett, D. (1997). *Cognition and Commitment in Hume's Philosophy*. Oxford University Press.
- Garrett, D. (2001a). Precis of Cognition and Commitment in Hume's Philosophy. *Philosophy and Phenomenological Research*, 62(1), 185–189.
- Garrett, D. (2001b). Replies. *Philosophy and Phenomenological Research*, 62(1), 205–215.
- Hume, D. ([1739–40] 2007). *A Treatise of Human Nature*, 2 vols. (D. F. Norton & M. J. Norton, Eds.). Oxford University Press.
- Hume, D. ([1751] 1998). *An Enquiry Concerning the Principles of Morals* (T. L. Beauchamp, Ed.). Oxford University Press.
- Hume, D. ([1777] 1985). *Essays Moral, Political and Literary* (E. F. Miller, Ed.). Liberty Fund.
- Joyce, R. (2009). Expressivism, Motivation Internalism, and Hume. In Pigden, C. R. (Ed.), *Hume on Motivation and Virtue* (pp. 30–56). Palgrave Macmillan.
- Kauppinen, A. (2010). What Makes a Sentiment Moral? In Shafer-Landau, R. (Ed.), *Oxford Studies in Metaethics*, vol. 5 (pp. 225–256). Oxford University Press.
- Korsgaard, C. M. (1999). The General Point of View: Love and Moral Approval in Hume's Ethics. *Hume Studies*, 25(1&2), 3–41.
- Köhler, S. (2013). Do Expressivists Have an Attitude Problem? *Ethics*, 123(3), 479–507.
- King, J. T. (1976). The Place of the Language of Morals in Hume's Second Enquiry. In Livingston, D., & King, J. T. (Eds.), *Hume: A Re-Evaluation*. Fordham University Press.
- Mackie, J. L. (1980). *Hume's Moral Theory*. Routledge.
- Magri, T. (1996). Natural Obligation and Normative Motivation in Hume's Treatise. *Hume Studies*, 22(2), 231–254.
- McDowell, J. (1998). *Mind, Value, and Reality*. Harvard University Press.
- Merli, D. (2008). Expressivism and the Limits of Moral Disagreement. *The Journal of Ethics*, 12(1), 25–55.
- Miller, A. (1998). Emotivism and the Verification Principle. *Proceedings of the Aristotelian Society, New Series*, 98, 103–124.

- Miller, A. (2013). *Contemporary Metaethics: An Introduction*, 2nd edition. Polity.
- Radcliffe, E. S. (2022). *Hume on the Nature of Morality* (Elements in Ethics). Cambridge University Press.
- Slote, M. (2010). *Moral Sentimentalism*. Oxford University Press.
- Smith, M. (2001). Some Not-Much-Discussed Problems for Non-Cognitivism in Ethics. *Ratio*, 14(2), 93–115.
- Stewart, C. (1976). The Moral Point of View. *Philosophy*, 51(196), 177–187.

About the Author



Shinya Aimatsu received his Ph.D. degree from the University of Tokyo in 2021. He is an assistant professor in the Department of Philosophy at the University of Tokyo.

✉ shinya.aimatsu@gmail.com

