

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

The Structure and Function of Emotion in Kant's Moral Theory

**Permalink**

<https://escholarship.org/uc/item/5rq8m2vg>

**Author**

DeWitt, Janelle A.

**Publication Date**

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Structure and Function of Emotion in Kant's Moral Theory

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Philosophy

by

Janelle A. DeWitt

2013

© Copyright by

Janelle A. DeWitt

2013

## ABSTRACT OF THE DISSERTATION

The Structure and Function of Emotion in Kant's Moral Theory

by

Janelle A. DeWitt

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2013

Professor Barbara Herman, Chair

The familiar terrain of Kant's account of the mind involves a two-fold distinction: between the two major faculties of cognition and desire, and between the higher and lower sub-faculties of each. But Kant's account of the mind contains a third major faculty that is missing from this picture—the faculty of feeling of pleasure and displeasure. Roughly, this is the faculty responsible for subjective sensation and feeling/emotion. However, at this point, a tension emerges. Just as the higher faculties of cognition and desire are identified with a function of reason, so, too, is the higher faculty of feeling. But if this is the case, does it mean that reason has or produces its own emotions? And if so, what could these emotions be like? Furthermore, wouldn't they conflict with the dichotomy thought to exist in Kant between reason and emotion (due in part to the metaphysical constraints of his theory)? In my

dissertation, I argue that Kant (implicitly) conceives of emotion in functional terms. That is, emotions are evaluative judgments that initiate action. These judgments can manifest themselves in a variety of forms, depending on the objects and principles involved. This turn to a functional conception thus gives Kant the flexibility to account for emotions across a wide spectrum, from the instinctual emotions of non-rational, embodied animals to the purely rational emotions of a non-embodied god. My dissertation first develops the basic functional structure of emotion implicit in Kant's work (found especially in the *Lectures on Metaphysics*). It then shows the implications this cognitive structure has for understanding several central features of Kant's practical theory, including the nature of animal versus human non-moral motivation, happiness, and finally, the moral feeling of respect.

The dissertation of Janelle A. DeWitt is approved.

Calvin G. Normore

John McCumber

Barbara Herman, Committee Chair

University of California, Los Angeles

2013

I dedicate this thesis to my father, Dennis—a man whose love for his children knew no bounds, and whose strength, compassion and deep sense of integrity was an inspiration to all who knew him.

The angels have indeed been with me...

## TABLE OF CONTENTS

Chapter 1	Feeling and Inclination: Rationalizing the Animal Within	1
I.	<i>The Standard Account of Feeling and Inclination</i>	4
II.	<i>The Possibility of a Cognitive Alternative</i>	9
III.	<i>The Cognitive Structure of Feeling and Inclination</i>	14
IV.	<i>The Rational Structure of Practical Experience</i>	23
Chapter 2	The Feeling of Happiness: An Expression of Our Humanity	34
I.	<i>Further Evidence of Kant's Cognitivism</i>	37
II.	<i>The Need for General Feelings of Happiness</i>	44
III.	<i>The Structure of Happiness</i>	48
IV.	<i>Sadness and the Passions</i>	59
Chapter 3	Respect for the Moral Law: the Emotional Side of Reason	65
I.	<i>The Faculty of Feeling of Pleasure and Displeasure</i>	68
II.	<i>The Lower and Higher Faculties of Feeling</i>	76
III.	<i>Respect</i>	91
Bibliography		101



## ACKNOWLEDGEMENTS

I would like to thank my entire committee for their support and often heroic efforts to help me through the process of writing my dissertation. But in particular, I would like to thank my chair, Barbara Herman, for helping me become the philosopher that I am, and for setting an example of the kind of philosopher and professor I hope to become.

I would also like to thank my family, Tina and Maria and Berneice DeWitt, for their incredible patience, love and support. Earning this degree would not have been possible without it.

## VITA

### EDUCATION

University of California Los Angeles, Los Angeles, California  
M.A. Philosophy

University of Chicago, Chicago, Illinois  
A.B. Philosophy in the Allied Field of Biology

### PUBLICATIONS

“Respect for the Moral Law: the Emotional Side of Reason,” forthcoming in *Philosophy*.

### SELECTED PRESENTATIONS

“Emotion, Motivation, and Kant’s Tri-partite Theory of the Mind,” Second Biennial  
Conference of the North American Kant Society, Cornell University, June 1, 2013

“Respect for the Moral Law: The Emotional Side of Reason,” Tennessee Values and  
Agency, University of Tennessee, November 2012  
Conference on Emotion, Catholic University of America, March 2012  
European Congress of Analytic Philosophy, Milan, Italy, September 2011

### SELECTED ACADEMIC/PROFESSIONAL HONORS

Dissertation Year Fellowship, 2010-2011  
University of California Los Angeles, Los Angeles, California

Dean’s Humanities Fellowship, 2008-2009  
University of California Los Angeles, Los Angeles, California

Graduate Research Mentorship Fellowship, 2007-2008  
University of California Los Angeles, Los Angeles, California

Graduate Summer Research Mentorship Fellowship, 2007  
University of California Los Angeles, Los Angeles, California

Dean’s Humanities Fellowship, 2004-2005  
University of California Los Angeles, Los Angeles, California

# Chapter 1

## Feeling and Inclination: Rationalizing the Animal Within

One bedrock feature of Kant's moral psychology is his description of humans as *sensibly-affected rational beings*. Implicit in this description is a duality of nature that gives rise to distinct forms of motivation—reason, stemming from our “higher”, rational nature, and inclination, stemming from our “lower”, animal/physical nature. However, what remains unclear from this division is how these two natures, and their relationship to each other, should be characterized.

Most commentators believe that Kant intends them to be understood as *entirely* distinct, in part because they stem from opposite sides of the metaphysical divide (the rational and the natural orders), and in part because reason, as the source of moral motivation, must remain unadulterated by its associated lower nature. Since these natures remain distinct, their characterizations are thought to proceed straightforwardly. Our higher nature, considered independently of its associated lower nature, shares the same basic characteristics as purely rational beings. The result is a nature entirely absent of any visceral experiences such as emotions, feelings, wants or needs. In contrast, our lower nature, considered independently of its associated higher nature, shares the same basic characteristics as non-human animals. The result is a nature entirely devoid of reason, where inclinations take the form of blind, impulsive, instinctual urges. The suggestion then is that these two natures could be separated, and little would change in how each, in itself, functions. In other words, implicit in Kant's theory is a strict dichotomy between reason and

emotion. As a result, the purpose of rational nature in human beings must be to govern their unruly, animalistic side. But because these natures remain distinct, it must do so from the outside—i.e., as a nature that stands *above* the animal as an independent ruler.<sup>1</sup>

This understanding of the structure of motivation is most clearly revealed in how many Kantians describe the incorporation thesis. A rough version of this account begins with the experience of pleasure in some object. This pleasure then elicits an unreflective attraction to the object in the form of a felt urge, or inclination. When the urge is strong enough, it immediately moves the animal to act. As a result, animals have only a mechanical form of agency. But humans, in contrast, have *free* choice, and somehow this must be accounted for. Since we share the same animal natures, the difference cannot lie in the form of their inclinations. Instead, it lies in the addition of a rational nature that can act as a gatekeeper. Inclinations must now first “propose” the action to reason for its approval. Reason, because of its standpoint above inclination, must then assess the proposal, and endorse (or reject) it before the inclination is allowed to move him to act. When the subject, *from the point of view of reason*, endorses the proposed action, he is said to “incorporate” it into his maxim. This incorporation thus provides him space between the inclination and the action for normative reflection. Without this normative distance, he would have no opportunity to exercise control over his agency, leaving him entirely subject to the mechanical forces of his animal nature.

---

<sup>1</sup> Tamar Schapiro provides a clear example of this view with respect to our lower nature when she states that “the non-rational form of agency exercised by our inclining part is structurally analogous to the type of agency that is exercised by nonhuman animals, creatures of instinct. The suggestion is that our capacity to incline has its source in our animal nature.” Because of this, she identifies the source of inclinations, i.e., our lower nature, as our “inner animal”. “Foregrounding Desire: A Defense of Kant’s Incorporation Thesis,” *Journal of Ethics* 15 (2011): 147-167.

Though the standard view is able to give an explanation for how reason can generate a form of agential control (via the incorporation thesis), it seems to come at the high price of a unified psychology. This view creates a deep alienation between fundamental parts of one's psychology—of who and what one is. Because this view holds that the subject's higher and lower natures are entirely distinct, he can identify with only one nature or the other at any given time. However, this creates a Dr. Jekyll/Mr. Hyde type persona, where the subject will be constantly flip-flopping between his two natures, never succeeding to identify with both from a single agential perspective. This is not to deny that there can, at times, be tension between elements of our psychology. But the view in question leads to a radical form of alienation—one in which the subject *never* experiences himself as unified. So unless he can take on the perspective of his rational *and* animal natures at the same time, he will continue to have a fractured psychology, and thus a fractured sense of what it is to be human.

Thus, in order to unify our psychology, and also to provide a more attractive theory of emotion and its role in the life of a human agent, we must recharacterize our two basic natures and the relationship that holds between them. But doing so will require rejecting the two related assumptions held by the standard view: that our two natures are entirely distinct, and that a strict dichotomy holds between reason and emotion. When we do, it then becomes possible for our rational nature in the practical domain to share the same characteristic function that it has in the theoretical domain—structuring and ordering input from our lower nature in a way that produces a unified experience of the world—but in this case, it will be a unified *practical* experience. In fact, I believe that our rational nature can

cross the metaphysical divide and structure our animal nature all the way down to sensations of pleasure and pain—not for theoretical cognition, but for practical cognition, i.e., for cognition “having to do only with the determining grounds of the will” (C2 5:20).<sup>2</sup> The result will be a sophisticated, *cognitive* account of non-moral motivation, one that can provide a framework to explain more complex features of Kant’s psychology, such as happiness and the passions (features that seem to require the involvement of reason in their very structure). When we see how deeply our rationality is infused into our lower nature, we then see that, on Kant’s account, even the animal within becomes rational.

## **I. The Standard Account of Feeling and Inclination**

Non-moral motivation, for Kant, is largely thought to be the result of an interplay between two basic elements of our psychology—feelings of pleasure and pain, and the desires they arouse. Since neither of these elements is thought to involve reason, human motivation will initially take the same basic form as that found in animals. To see how this interplay works, consider the act of drinking water. As a finite being, an animal will

---

<sup>2</sup> I have used two sets of translations from Cambridge University Press. The first is from the Cambridge Texts in the History of Philosophy: *Anthropology from a Pragmatic Point of View* (A), ed. and trans. Robert B. Loudon (2006); *Critique of Practical Reason* (C2), ed. and trans. Mary Gregor (1997); *Groundwork of the Metaphysics of Morals* (G), ed. and trans. Mary Gregor (1997); *Metaphysics of Morals* (MM), ed. and trans. Mary Gregor (1996); and *Religion within the Boundaries of Mere Reason* (R), ed. and trans. Allen Wood and George di Giovanni (1998). The second is from the Cambridge Edition of the Works of Immanuel Kant: *Critique of Pure Reason* (C1), ed. and trans. Paul Guyer and Allen Wood (1998); *Critique of the Power of Judgment* (C3), ed. and trans. Paul Guyer and trans. Eric Matthews (2000); *Lectures on Anthropology* (LA), ed. and trans. Allen Wood and Robert Loudon, trans. Robert Clewis and G. Felicitas Munzel (2012); *Lectures on Ethics* (LE), ed. and trans. Peter Heath and ed. J. B. Schneewind (1997); *Lectures on Logic* (LL), ed. and trans. J. Michael Young (1992); *Lectures on Metaphysics* (LM), ed. and trans. Karl Ameriks and Steve Naragon (1997); *Lectures on the Philosophical Doctrine of Religion* (LPDR) in *Religion and Rational Theology*, ed. and trans. Allen Wood and George di Giovanni (1996); and *Notes and Fragments* (NF), ed. and trans. Paul Guyer and trans. Curtis Bowman and Frederick Rauscher (2005). And finally *Idea for a Universal History with a Cosmopolitan Intent* (IUH) and *Speculative Beginning of Human History* (SBHH) in *Perpetual Peace and Other Essays* (Indianapolis: Hackett Publishing Co., 1983).

eventually come to feel a certain sort of pain we call thirst. Whether through hardwiring, following the lead of its parents, or trial and error, the animal soon discovers that drinking water alleviates this pain, resulting in a feeling of pleasure at having satisfied its thirst. Through mechanisms of association, the animal then begins to expect pleasure (or alleviation of pain) from drinking water. This expectation of future pleasure manifests itself in the present by affecting the animal's perception. When the animal is in a state of thirst, past associations of pleasure with water will make it appear more attractive or pleasant—i.e., as something “to be drunk”. The pull that this attractiveness exerts on the attention of the animal is thus felt as an unreflective urge to drink. Now, the more often the animal experiences pleasure from drinking, the more attractive water will appear to be, and the stronger and/or more consistent the urge will become—a reinforcing pattern that will eventually result in an inclination, or habitual desire, to drink water.

In the human case, little is thought to change. Because of past associations, whenever I am thirsty, I will likewise experience an urge for something to drink (usually coffee). But since I am also a rational creature in possession of a more advanced cognitive life, I can conceptualize the objects of pleasure and pain in ways that animals cannot. As a consequence, I might strongly associate pleasure with coffee generally (i.e., as *coffee*), but yet still associate displeasure with the particular coffee available on campus. I will then feel an urge to avoid any campus location and desire coffee from the nearby café instead. However, despite the involvement of reason in conceptualizing the object of pleasure, this object still functions no differently than in animals—as the *causal* source of the pleasure. The pleasure itself remains nothing more than a purely subjective sensation. So when it

elicits a desire (i.e., motivates), it does so simply because of the inherent pleasant (or painful) aspects of the sensation itself. In other words, Kant is thought to have held a non-cognitive, feeling based account of our emotions and appetites. And since, according to the picture above, there can be no desire without a feeling preceding it, these feelings take on a prominent role in non-moral motivation—i.e., they are the determining grounds of the will (MM 6:212, LM 29:894).

However, despite the widespread acceptance of this basic view of feeling, and the account of motivation that it gives rise to, few arguments have been offered in support of it. This is likely due, in large part, to the fact that Kant never explicitly offers a systematic account of feeling in any of his major texts, so there seems to be little material with which to work. Instead, the source for this view seems to be nothing more than his highly suggestive terminology. Considering that Kant not only describes them as feelings of *pleasure* and *displeasure*, but also appears to use the terms translated as ‘feeling’ (*Gefühl*) and ‘sensation’ (*Empfindung*) interchangeably, it would be natural to assume that a brute non-cognitive, or sensation theory is operating in the background. Non-cognitive views, as opposed to their cognitive counterparts, describe feelings in terms of their felt, experiential or affective quality. And as a simple view of emotion and appetite, it would lack enough sophistication to require much attention, thus explaining the absence of an explicit account in his work.

Additional support for the non-cognitive assumption is thought to be found in Kant’s generally negative attitude toward emotion. Whether it was due to his Pietist upbringing, or his Stoic influences, Kant has long been thought to have been, at best, highly skeptical of the value of emotion in human life, earning him the title of “cold-hearted” from almost the



beginning. But more often than not, he is portrayed as having been outright contemptuous of them. After all, they have the unwelcomed ability, through the interplay discussed earlier, to charm, distract, and distort our rational activity. In even darker moments, they are alien influences, rising up from our animal nature, that seek to overturn reason's authority by clouding the force of its pure moral law, all so that they can sow the seeds for vice-breeding passions. Because emotions have this tendency, Kant is thought to have insisted that reason must remain independent so that it can fully guard against their destructive influence. The consequence is a wholesale rejection of any theory in which emotion is even partially constituted by reason, because this would be tantamount to an invitation for the corruption of reason to its very core. A non-cognitive view, in contrast, would be a highly attractive alternative, because it would maintain a strict barrier between our higher and lower natures, and thus help to limit the negative influence of emotion on reason.

Now, though this general picture is a fairly plausible account of the kind of motivation and agency found in animals, in the human case, it turns out to be far less appealing. The source of its unattractiveness can be traced back to the brute non-cognitive view of feeling lying at its base. Not only is this theory weak in its own right, but it also gives shape to a deeply problematic structure of motivation and agency. Under this theory, emotions are reduced to the status of brute appetites and sensations, thus making compassion little different than thirst. Because thirst is a sensation resulting from blind, physiological processes, we view it as something that just happens to us. We never ask for a *reason* why someone feels (or fails to feel) thirsty, because it is not the sort of thing that reflects the character, values or beliefs of the person experiencing it. The same attitude will

now hold of compassion, fear, and anger. On this view, we cannot say that the reason we feel compassion is because we see someone in need, or feel fear because we believe an object to be dangerous, or feel angry because we are being subject to an injustice. Instead, like thirst, these will all be something we just happen to feel, explainable only in terms of our physiology. However, when feelings are understood in this way, their value in human life, moral or otherwise, becomes greatly diminished because they cannot be thought of as an expression of the character or commitments of the individual. Instead, their only value or purpose seems to be in helping to satisfy our most basic or instinctual needs for self- and species-preservation. It is not surprising, then, that this devaluation has been a deep and persistent source for criticism of Kant's entire practical project, and in particular, his theory of morality. Ever since the neo-Aristotelian turn, most modern moral theorists have come to recognize that emotions *are* a significant component of human life. Not only are they rich, complex phenomena that have a substantial impact on our practical engagement with the world, but as such, they are also thought to be genuine expressions of one's character. But to account for any of this more positive function of emotion, a cognitive theory is required; one that Kant's metaphysical framework seems to preclude from the start.

Unfortunately, the problem with this theory of feeling goes beyond relegating emotion to mere insignificance. Everything up to reason's approval at the point of incorporation can be characterized as a crude form of psychological hedonism, resulting in a merely mechanistic, instinctual form of agency. Specifically, because feelings motivate only in terms of their inherent pleasant or painful quality, all non-moral motivation is essentially a pursuit of pleasure and avoidance of pain. And since feelings, and the

inclinations they produce, are merely the effect of blind causal processes and brute associations, we essentially become a version of Pavlov's dog, responding unreflectively to whichever object triggers the strongest expectation of pleasure. So in order to get a genuine form of agency, and to moderate the hedonism embodied in our lower nature, reason must step in and somehow influence this process. But because the account of feeling at its base is non-cognitive, reason cannot affect the process itself (by helping to structure the feeling). Instead, it can only exercise control over its product, the inclinations, by approving or denying (via incorporation) the actions which they suggest. When reason is able to govern in this way, by having the final say before initiating action, it essentially becomes its own source of motivation. The problem with this, however, is that it does so as an alien force governing our lower nature from the outside. In other words, the solution to the problem of genuine agency becomes the source of the radical alienation between our two natures. So Kant's non-cognitivism leads us into an intractable dilemma—either we preserve our agency by fracturing our psychology, or we maintain a unified psychology by destroying our agency. The only way out is to replace the non-cognitive theory of feeling generating it, however unlikely this might seem to be.

## **II. The Possibility of a Cognitive Alternative**

At this point, however, I believe we should step back and ask, “Is this *really* Kant's view?” If we just look at its basic structure, it should actually strike us as being rather *un-*Kantian. In it, our two basic natures are characterized in a way that excludes any type of union or synthesis, resulting in a radical divide between them. And because they are also

both sources of motivation, they are now pit against each other in perpetual opposition. Our lower nature is worrisome because, as pure animality, it can be dangerous and unpredictable. Yet, reason has no control over it other than the ability to suppress through force.

The problem is, we find *nothing* like this on the theoretical side—not in how the function of reason is understood, nor in how it is thought to relate to our lower nature. In fact, what we see is the exact opposite. Rather than being the object of contempt, sensibility is regarded as essential to the cognitive life of humans because it provides the matter for cognition (as well as helps to rein in reason's tendency to overstep its proper bounds). Nor is reason set the task of merely filtering the raw input from sensibility, approving some of it and rejecting the rest. Instead, its essential function is to shape and organize this sensory input, and it does so in part by providing the very structure for the experience itself. The result is a single, highly conceptualized, and thereby *unified* experience of the world, one in which both the understanding *and* sensibility play crucial and harmonious roles.

So if *this* is how he conceives of the function of reason and its relation to sensibility on the theoretical side, why would the practical side be viewed any differently? Would we not expect instead that the same essential character of reason, and its relation to our lower nature, would be manifest in *both* its theoretical and practical forms? If so, then Kant would be more likely to value our lower nature as part of a practical life that is essentially human, because it provides the matter to the will required for us to engage in particular actions in the world. The task of practical reason would then be to govern, in the sense of unifying and harmonizing, by giving shape and structure to this matter. The result would be a single, unified, highly conceptualized and richly textured practical engagement with the world—

one in which our rational nature is expressed *through* its structuring of the emotions and appetites of our lower nature. In other words, given his broad theoretical commitments, we should actually expect a strongly *cognitive* account of feeling, one that allows our lower, animal nature to be transformed by reason, thus elevating it to the more dignified status of *humanity*.<sup>3</sup>

This expectation of cognitivism becomes even stronger when we closely examine certain passages in Kant's texts, such as his discussion of *sweet sorrow* in the *Anthropology*. This passage in particular raises questions for the standard, non-cognitive view of feeling and its role in motivation because it cannot explain the actions of the grieving widow that Kant mentions here. He states

...we also judge enjoyment and pain by a *higher* satisfaction or dissatisfaction within ourselves (namely moral): whether we ought to refuse them or give ourselves over to them....The object can be *disagreeable*<sup>4</sup>; but the *pain* concerning it *pleasing*. Therefore we have the expression *sweet sorrow*: for example, the sweet sorrow of a widow who has been left well off but does not want to allow herself to be comforted...(A 7:237).<sup>5</sup>

In this example, Kant is describing the sensible pain of grief as *pleasing*, and since it is pleasing, it motivates the widow to linger in her sorrow rather than be comforted by her good fortune. However, the contrast between the pain of grief and the pleasure of grieving

---

<sup>3</sup> "A human being has a duty to raise himself from the crude state of his nature, from his animality, more and more toward humanity, by which he alone is capable of setting himself ends" (MM 6:387).

<sup>4</sup> The German word here is *unangenehm*, which I believe is more properly translated (following Gregor) as *disagreeable* than (following Louden here) as *unpleasant*. To be disagreeable is to be a specific kind of unpleasant.

<sup>5</sup> In his discussion of sweet sorrow at C3 5:331, Kant further describes this higher satisfaction as a satisfaction (i.e., pleasure) that "rests on reason" because it "pleases merely in the judging" (as opposed to one that gratifies, or pleases in a sensation).

Kant draws here is in tension with the hedonic account of motivation discussed earlier. If the sensation of pain is itself motivating, then her grief would move her to be comforted. But the widow is not so moved. Instead, she lingers in the pain of her grief. But why? Kant's answer is that she finds the thought of doing so *pleasing*, despite there being no pleasant sensation involved. She does not linger because she feels a pleasant sensation *in* the painful experience of grief itself, if this is even coherent. (When one speaks in this way, the pleasure is not a sensation, but rather a form of assessment—i.e., the pain is something with which one “is pleased”.) Nor does she linger in her grief because she anticipates the action producing a sensation of pleasure. For example, she does not grieve in order to purge herself of the emotion, and thus alleviate the pain. And with the hedonic account, there are no other options—only a current or anticipated sensation of pleasure would motivate one to linger.

So what then are we to make of her finding grief pleasing? A straightforward interpretation suggests that she lingers because she sees there being something good or valuable to the painful experience of grief. Grieving is often a way of recognizing the value of the loved one who has died. And because she sees this action as being good or valuable, it is pleasing to her. This is what Kant means when he says that the higher pleasure (or satisfaction) *judges* the enjoyment or pain. Put another way, the hedonic account suggests that the positive assessment of a sensation of pleasure is part of the sensation's essence. Pleasure by definition attracts, pain repels. Kant denies this by distinguishing the assessment of the sensation (and its context) as a distinct type of pleasure (or displeasure). It is in this sense that we can have mixed states such as sweet sorrow and bitter joy (the

converse example), where a painful sensation can be pleasing (positively assessed and so maintained), and a pleasant sensation displeasing (negatively assessed and so avoided).

This second type of pleasure—specifically, the assessment of the sensation that causes the widow to linger—is, in essence, an action-initiating evaluative judgment, and as such, it explains her motivation without direct appeal to a sensation. Without any felt quality, however, it may seem odd for Kant to call the judgment itself a *pleasure*. But if we understand Kant’s use of *feeling of pleasure* to mean a positive, practical emotion (an ordinary sense of *Gefühl*), rather than a sensation, this is no longer the case. And though Kant’s language often falls short of explicitly referring to emotions as evaluative judgments (often taking the closely related form of *estimations* or *appraisals* instead), there are a few places where he does make such specific references. In his discussion on self-love in the *Lectures on Ethics*, he says “The love that takes pleasure in others is the judgment that we delight in their perfection” (LE 27:357/135). In his various notes and fragments, he states that “Feeling makes distinctions only for oneself; the judgment is not valid for others” (NF #1850 16:137/536). In the *Lectures on Anthropology*, he notes a distinction between pain and sadness. “Pain is mere sensation. Sadness [is] when I consider myself unfortunate. It is a judgment on the worth of my entire condition” (LA 25:1325/432).

If feelings are, for Kant, evaluative judgments, then we can begin to see why he elevates the *Faculty of Feeling of Pleasure and Displeasure* to one of major faculties of the mind on a par with the highly complex faculties of *Cognition* and *Desire*. (A full account of this tri-partite division will be given in the last chapter.) Following the standard account, his tri-partite division makes little sense. Feelings, understood as mere sensations, could never

merit this standing without severely undermining Kant's architectonic structure. But as evaluative judgments, or judgments about the good, the *Faculty of Feeling* becomes the interface between the other two faculties, i.e., between the world as it is (theoretical cognition), and the world as we want it to be (practical desire)—thus making it the faculty of practical cognition.

So there is clear evidence that Kant held some sort of cognitive view, but nothing in the passage discussed above tells us what particular form these feelings take. The challenge will thus be to uncover a structure of feeling that meets two basic demands. It must be able to account for the *full* range of human emotion discussed in Kant's texts—from basic, instinctual appetites and drives, to ordinary human emotions such as fear and compassion, to happiness, and even to the moral emotion of respect (happiness being the subject of chapter two and respect being the subject of chapter three). And it must do so while not violating Kant's metaphysical and anthropological framework.

### **III. The Cognitive Structure of Feeling and Inclination**

At this point, two general features of feeling are beginning to emerge. One, they involve some sort of evaluation or appraisal. And two, they are somehow the determining grounds of the will—i.e. they motivate action. But neither feature reveals much about the particular structure feelings must take. So the question still remains, what, exactly, are feelings? Kant asks roughly this same question in his *Lectures on the Philosophical Doctrine of Religion*. “But here I do not want to know *in what* I take pleasure, but rather *what* pleasure itself is.” He answers by stating “... pleasure itself does not consist in the



relation of my representations to their object; it consists rather in the relation of my representations to the subject, insofar as these representations determine the subject to actualize the object” (LPDR 28:1060, see also C2 5:21).<sup>6</sup> A similar description, in a somewhat more condensed form, is found in an early footnote of the second *Critique*. There he defines pleasure as “...*the representation of the agreement of an object or of an action with the subjective conditions of life,*” i.e., with the faculty of desire (C2 5:9fn). Considering our modern conception of pleasure as a type of non-representational or purely subjective sensation, this definition has tended to strike many as being quite cryptic. The non-cognitive assumption operating in the background puts the statement in such an odd context that it becomes difficult to make any sense of it at all. But if these passages are read in a straightforward way, Kant is clearly defining pleasure first and foremost *as a representation* (and this alone should have called the non-cognitive assumption into question). But more specifically, it is a representation, not of the object itself (a theoretical cognition), but of the object’s *agreement* with the subject’s faculty of desire—the source of his activity. And it is through this representation of agreement that we confer value onto objects, because “the

---

<sup>6</sup> Similar descriptions of pleasure as a *representation of an agreement* or *of a relation* can be found scattered throughout Kant’s writings. In his lectures, he says that “We have an inner principle for acting from representations, and that is life. Now if a representation harmonized with the entire power of the mind, with the principle of life, then that is *pleasure*. But if the representation is of the kind that resists the principle of life, then this relation of the conflict in us is *displeasure*. Objects are accordingly beautiful, ugly, etc., not in and for themselves, but rather in reference to living beings. *But what takes place only in reference to living beings, of that the ground must be in the living being*; accordingly there must be a faculty in the living being for perceiving such properties in objects. Pleasure and displeasure is thus a faculty of the agreement or the conflict of the principle of life with respect to certain representations or impressions of objects” (LM 28:247, see also MM 6:212, C3 5:209, LM 29:894, and NF #1021 15:457/408)

value of things always comes down to their concordance with subjects” (NF #715/495).<sup>7</sup> So feelings, in representing this relation, are the determinations of value, and it is in virtue of this that they motivate. In other words, they are action-initiating evaluative judgments. And it is with this understanding of feeling in mind that Kant says the faculty of desire is determined by a pleasure in an object.

Now, to see how this view of feeling works, consider the act of drinking coffee. That coffee is hot and caffeinated does not in itself motivate me to drink it. But when I recognize that I am currently cold, tired, and still have more writing to do, consideration of these properties *does* motivate me. The warmth and caffeine will alleviate my current discomfort and help me to continue working. Kant describes this relationship between the coffee and my subjective condition as a sort of “fit” that promotes my life or activity—in this case, my philosophical writing.<sup>8</sup> The representation of this fit is the pleasure that determines my will. Furthermore, when I judge the coffee in this way, as positively fitting (or *agreeing*) with my needs and activities, then I am at the same time judging the coffee to be *good* (i.e., to be *pleasing*). These pleasures, then, are a type of judgment that involves the

---

<sup>7</sup> Kant echoes this point in another fragment when he states “Things have many properties in themselves that remain even if they are not cognized by any rational being, but they never have any value (whether in sensation or appearance or concepts) except in relation to beings by which they can be cognized and for which they can be objects of choice” (NF #823/506). It can also be found in the *Groundwork*, where he says that subjective ends “are all only relative; for only their *mere relation to a specially constituted faculty of desire* on the part of the subject gives them their worth... (my italics)” (G 4:428).

<sup>8</sup> “What promotes our life, i.e., what brings our activity into play, as it were, pleases” (LL 45/31).

subjective/evaluative predicates of *good/bad*, or more specifically, of *agreeable/disagreeable* (LM 28:245).<sup>9</sup>

The same holds for displeasure. If instead I were ready for bed, then the caffeine in the coffee would hinder my ability to sleep (i.e., negatively fit with my needs/activities). I would then judge it to be *bad* (i.e., *displeasing*), and so would avoid drinking it. Any object that neither pleases nor displeases me in this way (neither promotes nor hinders my activity) is motivationally inert. It leaves me *indifferent* to it because it would have *no* connection to my will at all, the source of my activity (LM 28:253).

We now have the basic structure of feeling in hand. With these feelings, there is a sensible need, stemming from my animal nature, that drives the judgment—e.g., the alleviation of my fatigue. Certain objects are considered in relation to these needs, some of which are judged to fit. When there is a positive fit, the principle of self-love is determined to hold between the object and the subject. And since a need stemming from our finite, animal nature is the basis of this judgment, the resulting pleasure will determine *choice* (the *lower* faculty of desire) to make the object actual. Finally, when an object is determined to fit in this way, it is judged to be *agreeable* (LM 28:248).

This basic structure of feeling also determines *how* it motivates. That is, its motivational power stems from the fit it represents. I judge that the coffee fits with my needs because the caffeine it contains *will wake me up*. Implicit in this representation of fit, then, is the *expectation of a need being satisfied* by the coffee. It is this expectation that is

---

<sup>9</sup> The *beautiful/ugly* and the moral *good/evil* are also subjective predicates. But since the beautiful/ugly, as aesthetic pleasures, are not determining grounds of the will, I will not discuss them here. And since the good/evil are connected to *moral* motivation, they will be discussed in the third chapter.

the source of my *interest* in the coffee, and so is what ultimately motivates me to drink it.<sup>10</sup> Now, Kant distinguishes the actual satisfaction of my need, i.e., the alertness I come to feel, as a second type of pleasure—what he specifically refers to as sensations of *gratification/enjoyment* and *pain*.<sup>11</sup> These two types of pleasure (what I will call judgments and sensations of pleasures, respectively) must work together to motivate. That is, the judgments of pleasure can motivate only because of the expectation of gratification (satisfaction of a need) they represent. This follows because it is only through the gratification of a need that our animal life is furthered. The ability to gratify is thus the *condition* of the coffee being able to determine the will.<sup>12</sup> Because of this condition, I am said to drink it *for the sake of* the expected gratification. So when I judge the coffee as

---

<sup>10</sup> “The satisfaction that we combine with the representation of the existence of an object is called interest. Hence such a satisfaction always has at the same time a relation to the faculty of desire, either as its determining ground or else as necessarily interconnected with its determining ground” (C3 5:204).

<sup>11</sup> These are probably the closest things in Kant to the ordinary notion we have of “sensations of pleasure and pain,” though I believe he understands them more broadly. That is, gratification need not have any *particular* sensation-like phenomenology, though it often will (and so will be what provides the affective character of emotion). Instead, it only seems to require some sort of (empirical) conscious registration or awareness that an action was successful in satisfying a need, in whatever way that registration might manifest itself. Because of this, two very different experiences, the pleasant sensation in the taste of coffee or the psychological enjoyment of talking with an old friend, can both qualify as a form of gratification. And though we might not call the enjoyment felt in the latter case a *sensation*, on Kant’s account this is how it would be described. These sensations are the product of the interior sense (distinct from the inner sense of time) when it is affected by a representation from the faculty of cognition. Because of this, I believe Kant uses the term *Empfindung*, as opposed to the more general term *Gefühl*, in order to refer specifically to sensations of gratification/enjoyment and pain. As the product of a sense, gratification is thus an element of our sensibility (LM 29:1009, C3 5:205-7 and C3 5:331). From this, we see that, contrary to the non-cognitive assumption, even sensations of pleasure and pain for Kant have a complex structure involving reason. (Kant’s references to the interior sense are vague and infrequent, and he admits at one point to have not yet fully worked out the concept (LM 29:890). Even so, it still appears to play a significant role in his conception of pleasure. See section 15 of the *Anthropology* (7:153) for a very brief discussion.)

<sup>12</sup> Clearly, neither coffee nor its caffeine is agreeable to everyone. Kant notes this by pointing out that the expected gratification that serves as the basis for my judging the coffee as agreeable is dependent on the privately valid grounds of my own senses (LM 28:248-9 and C3 5:212). So this judgment holds *only for me*—i.e., it is *merely* subjective.

agreeable, I am judging it to be *good for* satisfying my needs (LM 28:252, LM 29:891, C2 5:59, C2 5:62 and C3 5:207). It is for this reason that Kant calls all empirical pleasures forms of self-love—because they are all ultimately directed towards satisfying one’s own needs.<sup>13</sup>

Kant’s account of feeling, in turn, gives shape to his account of desire. As noted earlier, motivation for Kant requires an interplay between pleasure and desire, but just not in the way the non-cognitive view suggested. On the default bi-partite conception of the mind most of operate with, where mental functions are divided between cognition and desire, desire tends to be thought of as an inherently intentional, or fully determinate state—e.g., hunger is characterized not as a blind urge, but as a desire *for food*. However, because Kant has a tri-partite theory of the mind, the form desire takes will be significantly different than that found on a bi-partite view. And this difference in form is what makes room for feeling to have the role it does as an evaluative judgment that determines the will.

In the second *Critique*, just before his definition of pleasure, Kant defines the faculty of desire as “a being’s *faculty to be by means of its representations the cause of the reality of the objects of these representations*” and shortly after, as “the faculty of the *causality of a*

---

<sup>13</sup> This understanding of the faculty of feeling of pleasure is not entirely unique in the history of philosophy. It bears strong similarities to the estimative power found in Avicenna and Aquinas, a power which is roughly responsible for perception/apprehension of the good. Simo Knuuttila notes that, for Avicenna, pleasure and joy are two emotions that belong to the *apprehensive* powers. Knuuttila’s description of pleasure in Avicenna bears a remarkable resemblance to the description of pleasure in Kant given above. “In *De medicinis cordialibus*, Avicenna says that sensitive pleasure is a perception of the fulfillment of a natural appetite or of the good functioning of the organism. This perception is pleasant—pleasure is the feeling aspect of the awareness of something positive taking place in the subject... Avicenna seems to assume that where there are desiderative emotions, the apprehensive acts activating them may already involve pleasant or unpleasant feelings” (223-4). The relationship between the apprehensive acts and the desiderative emotions they activate will also parallel Kant’s account of pleasure’s relation to desire (in the following discussion). For more on Avicenna and Aquinas’s account of the estimative power, see Knuuttila’s *Emotions in Ancient and Medieval Philosophy* (Oxford: Oxford University Press, 2004).

*representation with respect to the reality of its object*” (C2 5:9n, see also MM 6:211). Two things are notable about this definition. First, it is essentially defining the faculty of desire as the *source of our causality* for making the object of certain representations actual (LA 25:576-7/131). And second, it is not, itself, the source or cause of those representations. So their source must lie outside the faculty of desire, i.e., in the faculty of feeling. More specifically, the faculty of feeling must first represent an object to be good (i.e., judge it to fit with the needs or condition of the subject) before the faculty of desire can become the cause of the object represented. Therefore, the function of pleasure is to judge an object to be good in relation to the needs and circumstances of the subject, and the function of desire, in turn, is to be the cause of the object once it has.<sup>14</sup>

To see how this conception of desire works, we must begin with its source. Kant states that “...it is true of every created being that the desire for something always presupposes a need, and it is because of this need that I desire it. But why is this? *Simply because no creature is self-sufficient*, and so each one always has a need of many things” (LPDR 28:1062/398). All (sensible) desire thus originates in our particular subjective condition registered by the faculty of desire (the source of our life or activity)—i.e., with our particular needs, wants or activities due to our individual circumstances and constitution. In the example discussed above, it begins with my current condition of being cold, tired, and needing to continue working. But, desire only *begins* here, because merely feeling tired and

---

<sup>14</sup> “Pleasure is the representation of the agreement of an object with the productive power of the soul, and displeasure the opposite. The faculty of desire is the causality of the object which is produced. Accordingly, pleasure is the agreement and displeasure the conflict with our faculty of desire” (LM 29:894). Kant’s distinction between the faculty of feeling and the faculty of desire is what makes room for the aesthetic pleasures and for the pleasures responsible for mental acts such as faith (presupposing the postulates of pure practical reason). I discuss Kant’s tri-partite division more extensively in chapter three.

needing to wake up, or feeling cold and needing to warm up, cannot yet bring about any action. At this stage, these needs, or what we might call *proto-desires*, are more like action potentials that must first be engaged before they can move the subject to action.<sup>15</sup> They have this status because at this point they are indeterminate. That is, they have no intentional object yet towards which to direct their causality. To put this point another way, the faculty of desire is considered to be a source of causality because it can determine the *means* for making the object of a representation actual (C2 5:58-9). But these means cannot be determined until one has a representation of a particular object (e.g., coffee vs. tea) to be brought about. So in order for the latent causality of a proto-desire to be unleashed, something needs to provide it with an intentional object. In Kant’s terminology, the *elater* (proto-desire) needs an *elater animi*—i.e., something which can put “the ‘spring’ in the spring”.<sup>16</sup>

---

<sup>15</sup> I call these basic needs and wants proto-desires because Kant’s terminology is not always consistent. In the *Anthropology*, he seems to refer to them as undetermined desires, or peevish wishes. “The undetermined desire, in respect of the object, which only impels the subject to leave his present state without knowing what state he then wants to enter, can be called the peevish wish (one that nothing satisfies)” (A 7:251). In the *Religion*, he seems to describe them as propensities, or the “subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*)” (R 6:29). I believe it is also to these that Kant is referring in his description of concupiscence. “*Concupiscence* (lusting after something) must also be distinguished from desire itself, as a stimulus to determining desire. Concupiscence is always a sensible modification of the mind but one that has not yet become an act of the faculty of desire” (MM 6:213). The use of concupiscence in this passage has been misleading because of our contemporary association with sexual lust. But *concupiscentia* and *cupiditas* were two variations of Latin translation for the Greek *epithumia*—the first was used by Christian authors, the latter by non-Christian authors (Knuuttila, 156 fn153).

<sup>16</sup> “The causality of the representation with regard to (*later addition*: oneself is pleasure) the actuality of the object (*later addition*: objects in general) is desire (*later addition*: life; the *consensus* with life: pleasure). The representation must hereby, however, have a relation to the subject of determining it to action. This relation is pleasure, and indeed in the reality of the object, i.e., an interest (interest does not belong to judging). The interest rests on the satisfaction with our condition, which depends on the reality of the object. (That which carries an interest with it is called the *causa impulsiva*). An *elater* is the subjective receptivity to be moved to desire. ... An *elater* is the capacity of a *causa impulsiva* to determine the desire to a deed, insofar as it rests on

This is where judgments of feeling come in to the account of desire. When a feeling judges an object to be good for satisfying a particular need, it is, at the same time, setting that object as an end of the faculty of desire to be brought about (i.e., it is providing matter to the will). So through a determination by feeling, a proto-desire is transformed into a fully determinate desire that moves the subject to action. Such a determinate desire is thus a *Begierde*, or desire in the “narrow sense”—i.e., a desire that is “necessarily *preceded* by” a pleasure. When this interplay becomes habitual, it results in an inclination (MM 6:212).<sup>17</sup> So when my faculty of feeling judges coffee to be good for satisfying my particular needs, I then become *inclined to* the coffee—i.e., motivated to make the existence of coffee a reality. Feeling is thus the *elater animi* needed to activate the *elater*. In addition, because my desire is made determinate by feeling in this way, it will be an expression of the value I judge coffee to have given my current needs and circumstances. This overall conception of feeling and its relation to the faculty of desire is reflected in a dense passage of the second *Critique* where Kant states “...satisfaction with one’s whole existence...is a problem imposed upon him by his finite nature itself, because he is needy and this need is directed to the matter of his faculty of desire, that is, *something related to a subjective feeling of pleasure or displeasure underlying it by which is determined what he needs in order to be satisfied with his condition* (my italics)” (C2 5:25).

---

the constitution of the subject ” (NF #1021 15:457/408). He then describes the *elater animi* as “The possibility of a representation being a *causa impulsiva* (#1056 15:470-1)”, i.e., as a judgment of feeling.

<sup>17</sup> “As for practical pleasure, that determination of the faculty of desire which is caused and therefore necessarily *preceded* by such pleasure is called *desire* (*Begierde*) in the narrow sense; habitual desire (*Begierde*) is called *inclination*” (MM 6:212). Desire in the wide sense, or *Begehren*, suggests the general category which includes the full range of desire—proto-desires, fully determinate desires, wishes, etc.



#### IV. The Rational Structure of Practical Experience

One of the virtues of this basic structure of feeling and its relation to desire is that it can account for a wide range of appetite and emotion, from basic instinctual responses on the one end, to purely rational, moral emotions on the other, with ordinary human emotion falling in between. It can produce this wide array because it allows reason to have varying degrees of involvement, from ways of conceptualizing the subject and the object of the judgment, to providing different principles by which to judge the fit between the two. But even more importantly, when reason becomes involved in these ways, we begin to set our own ends rather than have nature set them for us. In other words, the more rational/cognitive a feeling becomes, the more *self*-determined, and so free, we will be.<sup>18</sup> Reason's involvement in feeling thus allows for a very different form of agential control to take shape. It no longer governs as an alien force, standing above our lower nature as an independent ruler. Instead, it now governs by guiding and directing feeling from within, shaping the very structure of our practical engagement with the world.

Now, to see how this works, consider the contrast between the simple forms of human feeling discussed earlier and our basic instinctual responses. To be a feeling of any kind, an object must be judged to fit with the subject (his condition) in a way that furthers the subject's life or activity. General categories of feeling are thus individuated on the basis

---

<sup>18</sup> "...[a human being] has a character, which he himself creates, in so far as he is capable of perfecting himself according to ends that he himself adopts. By means of this the human being, as an animal endowed with the *capacity of reason (animal rationabile)*, can make out of himself a *rational animal (animal rationale)*..." (A 7:321).

of variations within this basic structure. The primary point of departure between simple feelings and instinctual responses is in the two forms of need the subjective condition can take—propensity and instinct.

Most human feeling begins with a propensity. This might be best described as a state of *potential* fit with one or more objects.<sup>19</sup> For example, my being cold and tired is a state that can potentially fit with a slew of caffeinated beverages. But it can also fit with a nap, or even a bout of brisk exercise. Because this state potentially fits with a number of objects, it is thus indeterminate between them. So in order to act, I must first judge which of these objects will *best* satisfy my needs. I happen to like the taste of both coffee and tea, but when I am very tired, I will judge the higher caffeine content of coffee to be a better option. Yet when I am sick, I will judge tea to be better instead. And after several cups of either, I know to expect diminishing returns, so the alternatives to caffeine will then take on more value (in the form of a better fit with my need). Through this comparison of possible objects in relation to the current status of my needs, my actions will be highly adaptable and fine-tuned to the ever changing circumstances I find myself in. In other words, when I drink coffee, I am not acting unreflectively on an urge resulting from a physiological adaptation to caffeine. Instead, I am acting from a recognition of the value coffee has in relation to my current needs and activities, and the greater value it might have than sleep or exercise.

---

<sup>19</sup> “Propensity is the inner basis of all possible desires” (LA 25:580). “Propensity is distinct from actual desire. It is the possibility of desiring something, and is a predisposition of the subject to desire” (LA 25:796-7). “By *propensity* (*propensio*) I understand the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*), insofar as this possibility is contingent for humanity in general” (R 6:29\*). In this footnote, Kant describes propensity as the lowest (the most indeterminate in relation to an object) state of the faculty of desire. Above it is instinct, then inclination, and finally passion (a relationship to an object so strong that it excludes mastery over oneself).

Finally, when I judge coffee to be good in this way, it is still considered to be a general form of self-love because it is still directed at satisfying my own need. But, because it requires comparison among several potential objects, Kant calls it a form of *non-mechanical* self-love, or more specifically, a “self-love which is physical and yet *involves comparison* (for which reason is required)” (R 6:27). These propensities, and the judgments of feeling in relation to them, thus comprise the predisposition to *humanity*, a predisposition we have in virtue of being “a living and at the same time a *rational* being” (R 6:26).<sup>20</sup> It is thus a predisposition rooted in prudential reason—the form of reason subservient to our needs and natural ends (R 6:28).

However, in addition to these propensities, the subjective condition of *all* animal life must have another form, that of instinct. With simple feelings, the needs are indeterminate, often potentially fitting with several objects. So before one can act, reason (in the form of a judgment of feeling) first must select one of the objects to pursue, thereby making the proto-desire determinate. But if this was the only form that our subjective condition could take, then we would be in trouble. Until reason had gained some strength and had a fairly substantial amount of experience to draw from, the indeterminate nature of our propensities would make action impossible. A newborn infant could not judge that his mother’s breast milk would satisfy his hunger, and so would soon starve. As a consequence, nature must be able to act as a surrogate for reason, providing directly for the needs most essential to the

---

<sup>20</sup> The feelings discussed here are the most primitive form of humanity. As the comparison involved in the judgments become more complex, we will move from simple feelings in relation to particular needs to the more complex feelings of happiness and the passions. These are the two types of feeling/inclination that Kant was most concerned with in this passage of the *Religion*.

subject's survival. In other words, it must be able to "prejudge" a minimal set of objects that will reliably and consistently satisfy these needs for all members of its species. And it can do this because the sensibility of our lower/animal nature possesses what Kant calls an *analogue of reason*.<sup>21</sup> This analogue in turn has its own "principles", the innate rules of sensibility, which can generate what I call analogue judgments of feeling.<sup>22</sup> More specifically, through an innate principle of association, the animal is predetermined or hardwired to expect pleasure from certain objects, an expectation that then underlies a relationship of fit (an analogue judgment) without ever having experienced the object. So as it turns out, even instincts are determined by a "judgment" of feeling. It is just one that is innate and so makes our instincts determinate with respect to their objects from the moment of birth. This is how nature sets our ends for us.

When these analogue judgments determine the subject's essential needs, the result is a set of instincts, or natural drives responsible for self- and species-preservation (i.e., drives for nourishment, protection, and propagation). They function as natural drives because they have, in effect, an intentional object built into them by nature, and so are operational from the moment of birth (unlike propensity). This is how a baby can suckle without having been taught, why a chick will flee from a hawk even though it has never seen one before, and why

---

<sup>21</sup> "...for, if they are lacking consciousness, then they are also missing understanding and reason, and sensibility alone reigns. With animals one calls this an analogue of reason and there is an instinct of sensibility whereby they need no reason, but rather which an external being placed in them for acting, or for working according to instinct; the analogue of reason is the summation of all lower powers [i.e., of sensation, imagination, reproduction, and anticipation]" (LM 28:450/296, see also LM 28:594/354 and C3 5:464).

<sup>22</sup> To explain this in more detail would go well beyond the scope of this paper. But the rules of sensibility operative here are likely connected to the interior sense discussed in footnote 11, that is, "from the receptivity of the subject to be determined by certain ideas for the preservation or rejection of the condition of these ideas" (A 7:153).

even adult humans can sense beforehand whether something is fit to be eaten (LM 28:254, and SBHH 8:111).<sup>23</sup> But because the subject has no experience of the object yet, and so no particular representation of it, instincts are experienced as *blind* urges, i.e., as a “felt need to do or enjoy something of which we still do not have a concept” (R 6:29\*, see also LA 25:796-7 and LM 28:255-6). To put this another way, analogue judgments of feeling present themselves in a strongly perceptual way.<sup>24</sup> The hawk will just *look* dangerous to the chick, triggering an urge to flee even though the chick has no recognition (through experience) of the hawk as a predator. So instinct is a state of the faculty of desire that is more determinate than propensity, because it does have an intentional object. But it is not yet an inclination, because nature only provides it an obscure representation of the object, rather than the fully conscious or determinate representation that reason provides human inclination (R 6:29\*, LA 25:583-4).

Finally, the analogues of judgment responsible for instinct are also considered to be a general form of self-love, because they are still directed towards satisfying the needs of the subject. But because reason is *not* involved, Kant calls it a form of *physical*, or merely *mechanical* self-love, that is, “a love for which reason is not required” (R 6:26). These instincts, and the analogue judgments of feeling in relation to them, are thus the specific

---

<sup>23</sup> “Instinct—that *voice of God* that all animals obey—must alone have first guided the beginner. This permitted him to use several things for nourishment, but forbade the others (Gen. 3:2-3). However, it is not necessary to assume a special, but now lost, instinct for this purpose; it could have been met simply by the sense of smell and its connection with the organ of taste, the latter’s acknowledged sympathy with the organs of digestion, and, at the same time, the ability, which we are aware of even at present, to sense beforehand whether something is fit or unfit to be used for food” (SBHH 8:111).

<sup>24</sup> Kant believes that all sensible emotion can alter perception. Consider his discussion of hatred in the *Lectures on Anthropology*. “If one feels hatred, then the imagination shows everything from its most detestable side” (LA 25:1260).

elements of lower nature that comprise the predisposition to *animality*, a predisposition we have simply in virtue of being “a *living being*” (R 6:26).

Now, though nature can stand in as a surrogate for reason in animals and the very young, at least giving these creatures *some* form of agency, the type it can produce is very limited. Obviously, not all movement can be characterized as an *action*. A Venus Fly Trap is capable of movement towards an object, but because no representations are involved, it can only be characterized as a mechanical reaction to a physical stimulus according to the laws of nature. Plants, therefore, have no agency. But when an animal acts from instinct, it *does* act from a representation of an object as good or bad (an analogue judgment of an object as being fit or unfit to satisfy a need), so it does have at least a minimal form of agency. However, nature is limited from producing anything more robust. For genuine agency, the subject’s response to an object must be able to take into account his own particular needs and circumstances, because it is only in doing so that he can act with a conception of *his own* individual well-being in mind. And this is where nature fails the individual. When reason is not involved, the animal cannot conceptualize the object, its own needs, or the relationship between the two. So the subject’s need is in a fixed and relatively rigid relationship with the object set for it by nature. For the most part, a chick will always judge a hawk to be dangerous, whether or not the hawk is caged. Because of this, instinct only produces coarse-grained, stereotyped actions characteristic of all members of the species. And this is the problem. Because these are innate drives, nature must design them to provide for individual members of the species *on the whole*. That is, instincts are macro-level response mechanisms oriented towards the preservation and promotion of the *species*,

and it does so by providing for the *average* member of the species in the *most likely* circumstances. As a result, instincts can lead to unproductive and even highly detrimental behavior in some individuals, because they cannot take their *particular* circumstances into account. For example, a lame animal might be better off hiding from a predator, but its instinct will drive it to flee anyway. So at this point, nature can only provide for a form of agency that reflects the *general* needs and constitution of the average member of its species. In order to move closer to genuine agency, the subject's response to his environment needs to be adaptable, or capable of fine-tuning. Now, some non-rational animals are capable of moderating their instinctual responses through the associations they form from experience, allowing them to become a little more adapted to their particular environments. From repeated exposure, a campus squirrel's fear response to students can weaken. The more a dog participates on hunts, the more it will learn to expect patterns of behavior from a hare, and alter the manner of its chase (LM 29:949). But even so, the animal's judgment will still be between unconceptualized particulars, which then limit how it can respond. So without reason, animals can only expand the boundaries of instinct, never fully succeeding to step outside of them. They are thus limited to a mechanical form of agency.

Thus, in order for any sort of genuine agency to take shape, the subjective condition that action originates in must become open-ended, or indeterminate. Otherwise, the actions they produce cannot be closely tailored to individual—how he conceives is best for his own life to go given his unique needs, goals, talents, and circumstances. Kant describes this sense of freedom grounded in an indeterminate way of life in the *Speculative Beginning of Human History*.

[Through experimenting with objects he was not directed to by instinct] he discovered in himself an ability to choose his own way of life and thus not to be bound like other animals to only a single one...He stood as if at the edge of an abyss; for besides the particular objects of desire on which instinct had until now made him dependent, there opened up to him an infinitude of them, among which he could not choose, for he had no knowledge whatsoever to base choice on; and it was now equally impossible for him to turn back from his once tasted state of freedom to his former servitude (to the rule of the instincts). (SBHH 8:112)<sup>25</sup>

So until we can break free of the determinate life of instinct, where nature essentially structures our lives for us, there can be no room for *self*-determination, and so no freedom. This is where reason becomes crucial. It not only helps to make our original condition more indeterminate, but it also provides us the tools to then determine ourselves. It achieves both by helping the subject to produce a rich, conceptualized understanding of himself, his environment, and the relationship between the two.

Now, this conceptualization can vary in degree in each of the elements of the judgments of feeling, which then accounts for their cognitive range (from the non-rational/non-cognitive analogue judgments of instinct, to the empirical/cognitive judgments of ordinary emotion, to the purely rational/cognitive judgments of moral emotion). One of the first places this conceptualization has an impact is with the subject's representation of the object. The natural drive for sex, for example, is generally directed towards other members of the same species. But a man, as the object of this drive, can be thought of in more than one way. As the object of instinct, he will be judged to fit the needs of the subject

---

<sup>25</sup> Glimpses of this conception of freedom in an indeterminate form of life can be seen throughout Kant's work. "The characterization of the human being as a rational animal is already present in the form and organization of his *hand*, his *fingers* and *fingertips*... By this means nature has made the human being not suited for one way of manipulating things but undetermined for every way, consequently suited for the use of reason" (A 7:323).



*qua* member of the human species. When the object is represented in this way, the resulting instinctual drive is then relatively indiscriminate among members of the species. But when the subject can also represent him *qua* spouse, the relationship of fit will take on new complexity. As a result, the subject now has two ways to value her spouse, and so two ways in which she can determine her need (even with respect to the same object)—instinct and spousal love.<sup>26</sup>

The subject can also conceptualize himself and his own needs differently, both in a more specific and in a more general way. When he understands in greater detail the particular circumstances he finds himself in, and has a better understanding of himself and his own constitution, his relation to the object will again take on more complexity. Both a man and a woman have reason to fear a mugger holding a knife. But whether it is better to fight, flee, or give in will depend upon the particular constitution of each. The man might be timid and a very fast runner, making flight the best option. The woman, however, might be a little feisty and has significant self-defense training, making resistance the better course of action. Mere instinct, in contrast, elicits a one-size fits all, stereotyped response.

Seeing his needs and circumstances in a more general way also greatly expands the possible range of his responses beyond that of animals. He can come to see some of his own needs as shared, which then helps to bring him into community with others (a form of sympathy). He can consider his subjective condition in the context of societal norms,

---

<sup>26</sup> “...even if the inclination [of sex] to human beings fades away, not in so far as they are considered persons but merely as animal beings of the same species, in the inclination to sex, love to be sure can be passionate, but actually cannot be named a passion, because the latter presupposes maxims (not mere instinct) in proceedings with human beings” (A 7:270n40). This also allows our emotions to be reason responsive. If the subject values interaction with a man as a form of innocent fun, when she learns that he is married with kids, her desire for him will likely weaken, if not dissipate altogether.

leading to emotions of pride, shame, or resentment. He can consider his needs and their potential satisfaction in relation to others, which can form the basis for jealousy and envy. And finally, he can consider his needs collectively over a life-time, which allows for the emotion of happiness. (And as we will see in the next chapter, reason can also provide different principles of judgment in relation to these ways of understanding the subject, such as the principle of happiness.)

Thus, when reason is involved in this way, we begin to see a genuine form of agency take shape. The more richly textured the subject's conception of himself and his circumstances becomes, the more indeterminate and open-ended his subjective condition will be. This then makes room for him to act, not from nature's generic conception of his well-being, but from *his own* conception. In other words, through reason's involvement in feeling, he begins to set his own ends, by judging objects to be good in relation to his needs as he himself understands them. And this ability to *self-determine* is what makes him free.

So in the end, we see that, through the judgments of feeling, reason generates a complex evaluative framework that structures our practical point of view. That is, this faculty enables us to evaluate the world and our relation to it according to our current sensible needs, our long-term or overall well-being, and even morality, all from the *very same practical perspective*. I can judge that extra cup of coffee to be good for satisfying my current need, but at the same time as negatively affecting my health, and so also as potentially violating an indirect duty. The result is a unified deliberative field, and thus a unified psychology. And when coffee is judged in this complex manner, practical reason can then govern by giving shape and structure to the matter of my will. In other words, my

rational nature is expressed through the appetites and emotions of my lower nature, thus transforming my animality into humanity.<sup>27</sup>

Finally, we can now see how there can be varying degrees of freedom in relation to the varying degrees of reason's involvement in feeling. We are most free when we determine ourselves in a way that is *entirely* independent of our sensible needs, i.e., when we determine ourselves according to the moral law through the feeling of respect for it. And we are the least free when we are entirely determined by our sensibility, i.e., when nature determines our actions for us through the analogue judgments of instinct. However, in between morality and instinct, there is a wide range. The feelings discussed above constitute only the most minimal form. But as we will see in the next chapter, freedom from the natural drives of instinct reaches its maximum (short of morality) in happiness. By it we fully become the authors of our own actions, and so of our own lives.

*Nature has willed that man, entirely by himself, produce everything that goes beyond the mechanical organization of his animal existence and partake in no other happiness or perfection than what he himself, independently of instinct, can secure through his own reason. ... Specifically, he should not be led by instinct, nor be provided for and instructed by ready-made knowledge; instead, he should produce everything from himself. Provision for his diet, his clothing, his bodily safety and defense (for which he was given neither the bull's horns, the lion's claws, nor the dog's teeth, but only hands), all amusements that can make life pleasant, even his insight and prudence, indeed, the goodness of his will—all of these should be entirely of his own making. (IUH 8:19)*

---

<sup>27</sup> "This portrayal of mankind's earliest history reveals that its exit from that paradise that reason represents as the first dwelling place of its species was nothing but the transition from the raw state of a merely animal creature to humanity, from the harness of the instincts to the guidance of reason—in a word, from the guardianship of nature to the state of freedom" (SBHH 8:115).

## Chapter 2

### The Feeling of Happiness: An Expression of Our Humanity

Now, a rational being's consciousness of the agreeableness of life uninterruptedly accompanying his whole existence is *happiness*... (C2 5:25)

...all people have already, of themselves, the strongest and deepest inclination to happiness because it is just in this idea that all inclinations unite in one sum. (G 4:399)

**Toward practical philosophy.** The first and most important observation that a human being makes about himself is that, determined through nature, he is to be the author of his happiness and even of his own inclinations and aptitudes, which makes this happiness possible. He concludes from this that he has to order his actions not in accordance with instinct but in accordance with concepts of his happiness which he makes himself, [and] that the greatest concern which he has for himself here would be either that he forms his concept falsely or allows himself to deviate from it through animal sensibility, above all from a propensity to act habitually in opposition to his concepts. As a freely acting being, indeed in accordance with this independence and self-rule, he will thus have as his foremost object that his desires agree with one another and with his concept of happiness, and not with instincts; and the conduct befitting the freedom of a rational being consists in this form. First, his action will have to be arranged in accordance with the universal end of humanity in his own person and thus in accordance with concepts and not instincts, so that the latter will agree with one another because they agree with the universal, namely with nature. (NF #7199 19:272-3/462-3)

Without question, happiness is one of the most baffling features of Kant's practical philosophy. (And in this, there is stiff competition.) This is due, in large part, to the numerous and seemingly inconsistent ways in which he seems to describe it. On the one hand, happiness takes on characteristics of an emotion, and so is thought to belong to our lower, animal nature. In this form, happiness is a state of affairs in which one's life goes well (in the sense of having one's needs satisfied) (C2 5:124), coupled with a pleasant form

of awareness when it does (C2 5:22). And because we are natural beings, or beings with needs that must be satisfied in order to maintain our existence, this state is something that we all necessarily desire (G 4:415-16, R 6:6n), making happiness the object of an inclination (G 4:399), and so also a material determining ground of the faculty of desire (C2 5:25).

On the other hand, several descriptions of happiness seem to display features of rational activity. Considered in this sense, it appears to belong to our higher, rational nature. In this form, happiness is an empirical concept that we construct through the activities of reason and the imagination (C2 5:62 and G 4:418). And as a practical concept with normative force, it can act as a regulative principle of prudential reason (C1 A800/B828, C2 5:120), one whose task is to bring our inclinations into some kind of harmony (G 4:399 and R 6:58). Because of this, happiness is considered to be an end of the *will* (as the faculty of desire determined only by concepts) (MM 6:386-7). Finally, it is through these activities of reason that we are to become the authors of our own happiness—i.e., to give shape to our lives according to *our own* conception of well-being. Kant sees this authorship as an important aspect of happiness, because it is what allows us to break free from the chains of instinct (C1 A809/B837, NF #7199 19:272-3/462-3). In other words, it is one of two ways in which our freedom is expressed in the natural world.

Now, a fairly strong account of happiness could be made from either its characterization as a form of rational activity or as an emotion. The problem is, no *single* account taking in both of these characterizations seems possible, at least not without violating the metaphysical and anthropological framework lying at the foundation of his practical philosophy. Within this framework, there is thought to exist a deep division

between the rational and the natural orders. As a consequence, our higher and lower natures are assumed to be entirely distinct, which in turn gives rise to a radical dichotomy between reason and emotion. In other words, our higher, rational nature is entirely devoid of any visceral experiences such as wants, needs, urges, or feelings, and our lower, animal nature is entirely devoid of reason. And because of this framework, it is thought that Kant must have adopted a brute, non-cognitive theory of emotion, one in which emotions are considered to be pleasure and pain-like sensations that blindly impel the subject to action. However, when our two natures and the relationship between them are understood in this way, the two conceptions of happiness, as rational activity and as an emotion, become deeply incompatible. The tension between them can surface in any number of places. For example, if inclinations are the result of blind physiological processes, and so not a product of rational activity, then how can we have an inclination to happiness and still consider ourselves to be its *author*? Or, if an inclination cannot, by its very nature, have a concept of reason as its object (in anything other than a causal sense), then how can happiness be both an empirical concept of reason and an object of inclination? In other words, because the two conceptions of happiness tend to fall on opposite sides of the metaphysical divide, any attempt to reconcile them will inevitably conflict with his deeply situated framework.

This leaves someone sympathetic to Kant with a very limited set of options. Either he can accept that two competing conceptions are operative, and so attempt to minimize the conflict between them. Or he can take sides, and try to explain away the passages supporting the other conception. Neither of these options, however, seems very palatable, in large part because I believe that Kant was neither confused nor misleading on the issue. So I

suggest that we opt for a more radical strategy instead, one in which we reject the assumption of a dichotomy between reason and emotion, and the non-cognitive theory that it entails. In its place, I suggest that we develop the cognitive account of emotion implicit in his work, one that will provide the theoretical resources necessary in order to give a single, unified account of happiness. More specifically, a cognitive theory will allow a human being to be the “author of his happiness and even of his own inclinations”, because it will then be possible for reason to be involved *in the very structure of the inclination itself*. When it does, his conception of his own happiness will become the evaluative framework that structures his non-moral practical point of view. He will then be able to set his own ends (according to this conception of happiness), instead of having nature set them for him, making his life *his own* for the first time. For this reason, happiness is the concept central to Kant’s account of non-moral motivation, because it is by the subject’s own conception of it that he can finally throw off the shackles of instinct and exercise the freedom that his own humanity affords him.

## **I. Further Evidence of Kant’s Cognitivism**

As we saw in the last chapter, the standard account considers non-moral motivation to be the result of an interplay between two non-cognitive elements of our psychology—feelings of pleasure and pain, and the unreflective desires they arouse. The story begins with some pleasurable or painful sensation experienced in connection with an object—e.g., the painful feeling of the rapidity of my heartbeat and the trembling of my limbs when I am confronted by a mugger. These feelings then elicit a desire to flee the location, or fight if I

cannot. If I am approached again in the same area, my association of fear with that location will be reinforced, and as a result, I will feel a strong urge to avoid it. Our experience of our non-moral agency is thus one of being pushed and pulled by these blind forces.

However, I also argued that this view of feeling is not Kant's own. And because it is not, its incorporation into Kant's practical theory has generated numerous problems and/or interpretational difficulties—including a radical form of alienation between our higher and lower natures. One of the reasons for these problems is non-cognitivism's failure to capture the complexity of emotion evident in Kant's discussions, a complexity that is required in order to account for the more complicated elements of his psychology, such as happiness, the passions and the moral emotions. This need for a more sophisticated theory of emotion thus indicates that Kant is more likely to have adopted a cognitive theory instead.

I then showed that certain passages, such as his discussion of sweet sorrow, back up this claim, because they bear the general marks of a cognitive view. Central to the general cognitive thesis is the claim that emotions are appraisals or evaluations of the significance or worth of an object, circumstance or event. Thus, on this view, what is essential to the emotion of *grief*, for example, is not the feeling of emptiness or muted perception of the world resulting from the loss, but the recognition of the significance of the loved one who has died. Likewise, what is essential to *fear* is not the feeling of trembling, but the judgment that something in one's environment poses a threat to one's life or well-being. And when the objects are judged in these ways, they take on a certain perceptual or deliberative salience that, in turn, influences action, such as lingering in one's grief or fleeing from a mugger.



A second mark of cognitivism, not discussed earlier, is the presence of detailed conceptual analyses of various clusters of closely related emotions, such as fear, anxiety and dread, because such analyses are possible only in terms of the cognitive component involved. In this regard, non-cognitivism is severely limited because it can only individuate on the basis of qualitative differences in feel, and variations on these qualities quickly run out or begin to blur together. Consequently, fear, anxiety and dread are thought to collapse into roughly the same emotion because they all *feel* much the same way. But cognitivism can make sharp distinctions between them. All three share the basic form of fear—an assessment of danger. But fear proper can be distinguished from anxiety and dread because its object is present or immediate harm. Anxiety and dread, in contrast, involve future harm. But the harm in anxiety is merely possible or indeterminate, whereas that of dread is known and impending. Distinctions between emotions can thus be made according to even the most subtle variation in the underlying evaluative judgment.

What may be surprising to many is that Kant's lectures are filled with this kind of detailed conceptual analysis. But perhaps more importantly for present purposes, these discussions also show the incredible level of complexity that emotion can take. His discussion of the cluster of jealousy, envy, and *Schadenfreude* found in the *Lectures on Ethics* is one such example (a set of distinctions particularly difficult for non-cognitivism because they do not have much affective character among which to make distinctions). Here, Kant notes that human beings have a natural tendency to estimate our own worth (stemming from a legitimate pretension to equality (LA 25:609 and R 6:27)). But such estimations require comparison, either with an absolute standard, such as the idea of

perfection (a form of *self-esteem*), or with a relative standard, where I compare myself to others (a form of *self-love*).<sup>28</sup> When I estimate my worth in relation to the absolute standard and fall short (a negative estimation), the only way to alleviate the resulting pain of *self-contempt* is to perfect myself in relation to this standard. But since this is often difficult, I am prone to estimate my self-worth by the weaker, relative standard instead. With this standard, I am much more likely to come out with a higher, inflated, and therefore often false estimation (NF 2:312/13), since most of us (as the relative class) fall far short from the ideal. In other words, though I may look like a scoundrel when compared to Mother Theresa, I will look like a saint next to Hitler.<sup>29</sup> The problem with the relative standard, however, is that it is *relative*. So when I estimate my worth in relation to this standard and fall short, I now have two ways to alleviate the resulting pain of *self-hatred*. I can either increase my own worth, or take the easier route and destroy the worth of others. With the latter option, my self-hatred is then redirected into a hatred of others (i.e., an assessment of the other as being unworthy). This general feeling of hatred thus becomes the source of all vice-breeding passions, including those of jealousy, envy and *Schadenfreude* (LE 27:436-44, LE 27:691-8, and MM 6:458).

---

<sup>28</sup> “Self-love differs from esteem.... We esteem what has inner worth, and love what has worth in a relative sense” (LE 27:357). The difference in the standards of comparison when we estimate our own worth is thus what distinguishes self-love/self-hate from self-esteem/self-contempt. More specifically, judgments against an absolute standard determine one’s true, inner worth, while judgments against a relative standard only determine relative worth.

<sup>29</sup> “Men have two means of estimating themselves; when they compare themselves with the Idea of perfection, and when they do in relation to others. If we estimate ourselves by the Idea of perfection, we have a good standard of measurement; but if we do so in comparison with others, we may often come to the very opposite conclusion, for now it is a matter of what sort of people they are, with whom we compare ourselves. ...but if he compares himself to others, he may still have great worth, in that those with whom he matches himself may be great rogues.” (LE 27:436, see also LE 27:359).

This [tendency to compare] is the origin of jealousy. When men compare themselves with others, and find these perfections, they become jealous of every perfection they perceive in the other, and try to diminish it, so that their own may stand out the more. This is disparaging jealousy. But if I try to add to my perfections, making them equal to the other's, this is emulating jealousy (LE 27:437).

Later he states

We grudge [are jealous] when displeased at another's advantage; we are too much put down by his good fortune, and therefore grudge it to him. But if we are displeased at the fact that the other has any share of happiness, that is envy. So envy is when we wish imperfection and ill-fortune to others, not so that we might ourselves be perfect or fortunate in consequence, but so that in that case we might alone be perfect and fortunate. The envious man wishes to be happy when all around him are unhappy, and seeks the sweetness of happiness in this, that he alone enjoys it, and all others are unhappy (LE 27:438).

Envy extends not only to all the praiseworthy features necessary to man, but also to talents and fortunate circumstances of any kind.

If the other's advantages arouse merely distress in a man, because on comparing his worth with the moral standing of the other he feels himself degraded, this is merely *misliking*, or *invidia in genere* (envy without ill-will). He feels merely his own unworthiness by the comparison made, and is discontented because the other possesses advantages. But such envy becomes *invidia qualificata* (malicious envy, or spite), i.e., *livor*, when within him there is simultaneously awakened the desire to lessen those advantages, and to injure the other on that account (LE 27:694).

Jealousy, then, is a feeling of discontent when another possesses what I also want (e.g., happiness, advantage or good fortune), but do not have. If it remains a form of *misliking*, i.e., grudging without hatred or ill-will toward the other, then Kant characterizes it as *emulating* jealousy, where I am motivated to improve upon my situation in order to bring myself into equality with others. When considered in this regard, jealousy is not wholly bad. But the problem is, it can easily come to involve *hatred* or ill-will, and when it does, it turns into envy, or *disparaging* jealousy. Envy, then, is distinct from mere jealousy in that it

seeks the destruction of the other's happiness or advantage in order to raise the estimate of one's own in relation. But what makes this passion so dangerous (as if its destructive tendencies were not bad enough) is that it will not stop until the subject *alone* is happy. In other words, because his own happiness is heightened relative to the misfortune of others, it will reach its maximum only when those around him are miserable. The only way to protect against this vice is through *self-esteem* rather than *self-love*, because the standard at the basis of self-esteem is absolute. Since this standard cannot be diminished, the only way to improve one's estimation is through self-perfection, thus preventing the destructive forms of jealousy from taking root.

*Schadenfreude* is similar to both jealousy and envy, because each of them involves taking a certain sort of pleasure in the misery of others. But unlike jealousy and envy, which take pleasure in this misfortune only *indirectly*, as a means to heighten one's own happiness, *Schadenfreude* takes pleasure in the misery of others *directly* (LE 27:440). In other words, he is the extreme opposite of the naturally sympathetic man in the *Groundwork* (G 4:398, see also C2 5:34). Just as the sympathetic man has a psychological need to spread joy around him, and takes immediate pleasure in doing so, the man of *Schadenfreude* has a psychological need to spread misery. "Since the man given to *Schadenfreude* does not remain a mere spectator of the other's sufferings, but at least participates in them through the glee that he feels at the spectacle, this vice, as an evil inclination, already seems to be the farthest removed from humanity" (LE 27:697). In other words, *Schadenfreude* is considered to be a form of *inhumanity*, and so is the worst of the three (LE 27:693).

Since the basis of the passions is a form of self-love, or estimation of self-worth using the relative standard, we can now see why Kant states that animals are not subject to the passions (A 7:266 and A 7:269). Because the passions involve evaluation based on comparison, they *require* the involvement of reason. Kant makes this claim explicitly in the *Religion*, where he discusses the basis for his distinction between the form of self-love shared with animals, which he calls a predisposition to *animality*, and the form only found in humans, which he calls a predisposition to *humanity*. The first, as a form of *mechanical* self-love, is “a love for which reason is not required”, and so is the basis of the instinctual responses we share with animals (e.g., the drives for self-preservation and propagation of the species). The latter form of self-love is “physical and yet involves comparison (for which reason is required); that is, only in comparison with others does one judge oneself happy or unhappy” (R 6:26-27). In other words, all forms of non-mechanical self-love involve, *as part of their definition*, an evaluation of oneself in relation to others (or to a standard) which requires reason. So it is only out of the form of self-love stemming from our predisposition to humanity that the passions can arise. In other words, only this form of self-love produces an inclination to equal worth, an inclination which can quickly transform into an inclination to gain worth in the opinion of others, and continue to descend into a desire to acquire superiority for oneself over others. “[T]he vices that are grafted upon this inclination can also be named the vices of *culture*, and in their extreme degree of malignancy (where they are simply the idea of a maximum of evil that surpasses humanity), e.g., in *envy*, *ingratitude*, *joy in others’ misfortunes*, etc., they are called *diabolical vices*” (R 6:27). This is why Kant

calls *Schadenfreude* a form of inhumanity, because it is a perversion of the very humanity that lies at its source.

Though Kant does not explicitly assert a theory of emotion in any of these discussions of the passions, it would be very difficult indeed for non-cognitivism to explain them. Because the passions are defined in terms of comparative estimations or evaluations that require reason, an explanation only seems possible on a cognitive theory. In the previous chapter, I presented an outline of the cognitive structure I believe Kant adopted. But we now must build upon this structure, because we see from the discussion above that at least two more elements are needed. First, different standards of evaluation, or principles of judgment, must be possible. And second, feelings must be capable of comparison, not just among objects (as discussed in the last chapter), but in relation to these standards. Both of these, as it will turn out, are also needed to account for happiness.

## **II. The Need for General Feelings of Happiness**

In the last chapter, I developed the basic cognitive structure of emotion implicit in Kant's practical theory. As noted previously, all emotion shares this same basic structure. But the variations possible among its elements allow Kant to generate several different categories of emotion, each with its own character and unique function. An account of happiness will thus build off of this basic structure. To quickly review, with any type of feeling for Kant, an object must be judged to fit with a subject in a way that either furthers (or hinders) the subject's life or activity. If the object furthers life, then it (as the object of the pleasure) determines the will. Consider the feeling of fear discussed earlier. That

someone is carrying a gun does not, in itself, motivate me to avoid him. But when I take note of the fact that I am also walking down a dark street late at night, I *will* likely change direction. A potential mugger carrying a gun would put my life at risk (i.e., negatively fit or disagree with my needs and circumstances). As a result, I would judge the stranger to be *dangerous* (i.e., *bad* or *displeasing*), and would immediately take steps to protect myself from him. In contrast, a police officer would provide a degree of security (i.e., positively fit/agree), and so would be judged to be *protective* (i.e., *good* or *pleasing*) instead. In that case, I would simply continue on as before.

From this example, we also see the relationship between the elements of the judgment. With any feeling, there is a sensible need (or proto-desire) stemming from my animal nature that drives the judgment—e.g., my safety.<sup>30</sup> Certain objects are then considered in relation to these needs, some of which are judged to agree/disagree. When there is either a positive or negative fit, the principle of self-love is determined to hold between the object and the subject. And since a need stemming from our finite, animal nature is the basis of this judgment, the resulting pleasure will determine *choice* (the *lower* faculty of desire). When it does, the original need becomes determined to this object, thus inclining or motivating me to make the particular object actual.

This basic structure of feeling also determines how it motivates. That is, its motivational power stems from the fit it represents. I judge that the stranger disagrees with my needs because the gun he is carrying could cause severe physical harm. Implicit in this

---

<sup>30</sup> This is true only of the feelings of our lower nature. As I will show in the next chapter, there are feelings of our higher, rational nature that are not based on any type of need.

representation of negative fit, then, is the expectation of a sensation of pain—i.e., of a need being thwarted or even intensified by the stranger. This expectation of pain is thus the source of my *interest* in the stranger, and so is what ultimately motivates me to avoid him. So these particular judgments of feeling can motivate only because of the expectation of gratification/pain (satisfaction/dissatisfaction of a need) they represent. This follows because it is only through the thwarting of a need that my animal life is hindered. The ability to pain is thus the *condition* of the stranger being able to determine the faculty of desire. Because of this condition, I am said to avoid him *for the sake of* the expected pain. So when I judge him to be disagreeable, I am judging him to be *bad for* satisfying my needs (LM 28:252, LM 29:891, C2 5:59, C2 5:62 and C3 5:207). As such, the stranger is only bad conditionally (instead of absolutely, as in the case of moral evil). And since these pleasures and displeasures all aim at satisfying my needs, Kant calls these feelings a form of self-love.

So far, I have shown how judgments of feeling can choose which among a set of objects will best satisfy a particular need. But because we are finite beings with limited resources, this is not yet enough. We need to be able to determine not just which objects will satisfy our needs, but even which needs to satisfy at all. Kant makes this point in a passage of the *Lectures on the Philosophical Doctrine of Religion*.

Thus the Stoics thought the ideal of the sage, as one who would feel no compassion for distress, but would feel no greater delight in anything than in remedying all distress. This ideal is not possible for human beings; for an incentive must be added to my cognition of the good before I can actually will to produce the good. This is because my activity is limited, and thus if I am to apply my powers to the production of some good I must first judge whether in this way I am not using up resources which might have produced some other good. Therefore I need certain incentives to determine my powers to this or that good, since I do not have resources sufficient for the actual production of everything I cognize to be good. -- Now these incentives



consist in certain subjective relations which must determine my being well-pleased in choosing, subsequent to the determination of my *well-pleasedness* in judging or my cognition of the good. If this subjective relation were taken away, then my choice of the good would be removed. (LPDR 28:1065-6)

Kant notes here that, because we are finite beings, our needs often eclipse our means to satisfy them. It is inevitable, then, that some needs will be satisfied at the expense others. Now, if there were no mechanism in place to coordinate and prioritize these needs, then we would be left entirely hostage to the momentary fluctuations of our current subjective condition (even on a cognitive view of feeling). In other words, our actions would be entirely determined by the strongest, or most immediate need felt at the moment, regardless of the impact satisfying that need might have in relation to the possible satisfaction of any other. A person suffering from gout, for example, will have no resources to resist his current desire for a rich meal, a desire that if satisfied will likely lead to a great deal of pain in the future. As rational beings, however, we cannot simply be left subject to the contingent and ever changing tidal flows of our current condition. Somehow reason must provide us a way to act from a conception of our long-term, overall well-being, and the value particular actions have in relation to that conception. According to Kant, it does so by providing us concepts of reason, such as happiness, by which we can organize these needs, thereby bringing about a balance or harmony among them.

But the question is—how does this happen? In the passage of the *Groundwork* related to gout, Kant describes the man as having two different, and competing, inclinations—a single inclination, for what I am guessing is a rich meal at dinner, and a general inclination to happiness. Implied in this discussion is that the man *ought* to act on

his inclination to happiness, but often does not. That is, Kant means this to be an example of imprudent action, and explains it by pointing out that the concept of happiness is indeterminate and so lacking some degree of normative force. But this example tells us something significant about the nature of an inclination to happiness. If it was simply one among many similar inclinations, then Kant could not make this claim. The desire for a rich meal and the desire for health would be pushpins and poetry at this point. So an inclination to happiness must be different in kind from the particular inclination. And this difference lies in the fact that it reflects a different principle of evaluation, one that will make the resulting inclination more objective—the concept of happiness. Any such difference in inclination, however, must come from the judgment of feeling that gives rise to it. So it must be possible to incorporate the principle of happiness into the structure of a judgment of feeling. The result will be a set of feelings that can order and rank our needs, and when they determine the will, the inclinations they produce will reflect this harmony.

### **III. The Structure of Happiness**

The essence of happiness is thus a harmony among the needs of our subjective condition. But how, then, can the structure of feeling produce it? And how is the concept of happiness involved in this structure? We saw earlier that, with any type of feeling for Kant, an object must be judged to fit with a subject in a way that furthers the subject's life or activity. Within this basic form, certain variations are possible. The object can be anything from particular objects like coffee, to means such as money, to concepts such as friendship and health. The principle can be one of several forms of the principle of self-love

(considered generally as a principle directed toward satisfying the needs of our animal nature). So far we have uncovered the particular principle of self-love (and its analogue in sensibility), and the possibility of a principle of happiness or general self-love.<sup>31</sup> Finally, the subjective condition can be understood in one of two ways—either in terms of a particular need or in terms of a sum total of needs. The particular feelings of self-love discussed in the last chapter, because they are *particular*, are judgments between a single, concrete object or state of affairs and a particular need of the subject according to the particular principle of self-love. Because of this, when they determine the faculty of desire, they can bring about a particular action *in the world*. That is, when I judge a cup of coffee to be good for satisfying my current need to wake up, I will be motivated *to head to my local café*. In contrast, when I judge that exercise is good for satisfying my need to be healthy, I will merely be motivated to exercise in general, i.e., to adopt exercising as a general maxim upon which to act. However, because I am a physical being, for this motivation to lead to *action*, the object of the judgment must be concrete or particular instead of abstract—e.g., to go for a run on the trail nearby. For this reason, particular feelings are important to our agency, because without them, we would be unable to physically act at all.<sup>32</sup>

---

<sup>31</sup> The final principle of our lower nature (not addressed in this dissertation) will be *universal* or *rationalized* self-love, a form of self-love that involves the concept of *universal happiness*, or the happiness of *all* rational beings considered in totality (C2 5:36 and R 6:45†). The incorporation of this principle into feeling is what helps us to shape our lower nature to fit better with morality, and so to move from virtue (a state of conflict between our animal and moral natures) towards holiness (a state of harmony between our two natures where the moral law no longer needs to *command*) (C2 5:32, C2 5:34, and C2 5:84).

<sup>32</sup> In this sense, they become the practical analogue to intuition. This can help to explain why Kant often couples the terms, as he does, e.g., in the *Groundwork* when discussing the three ways of representing the principle of morality. “However, there is yet a dissimilarity among them, which is indeed subjectively rather than objectively practical, namely to bring an idea of reason closer to intuition (according to a certain analogy) and thereby to feeling” (G 4:436). Bringing it closer to feeling in this sense is to particularize the moral law by

Feelings of general self-love, in contrast, will have a very different shape, one that in turn will help to explain their harmonizing function. The primary point of departure from particular feelings is in how the subject, or his condition, is to be understood. All feelings of the lower nature are concerned with furthering the animal life of the subject—i.e., promoting the activity of his lower nature directed towards maintaining his physical/sensible well-being. Furthering this life is thus understood as satisfying the sensible needs stemming from his finite nature. With particular feelings, as the name suggests, this activity is considered in terms of a particular need. When this need is satisfied by an object, it thus follows that only one aspect of his animal life is furthered. But the problem is, as noted in the previous section, the satisfaction of a single need will not guarantee an increased satisfaction of the whole. A rich meal for a man with gout may further one aspect of this activity (by satisfying a need for something tasty), while also hindering several others (by harming his health, depleting his bank account, etc.). The reason for this is that, as a *particular* feeling, it considers the need of the subject in isolation.

Thus, in order to provide for the overall or long-term well-being of the subject, feeling must be able to take a *general* form, i.e., a form which considers the satisfaction of a particular need *in relation to the subject's totality of needs*. In essence, these feelings must be capable of making a *comparative* judgment. Now, when the subject's needs are considered in this way, a different relationship of fit will result. That is, an object will fit with the subject only if it positively contributes to a maximal satisfaction of his total set of

---

imposing its form onto a particular feeling so that the subject can more easily *act* from it. How the moral law can do this, in the form of respect, is addressed in chapter three.

needs. Implicit in this representation of fit, then, is what we might call the *form* of the life or activity of the animal—i.e., how each of the particular needs of the subject relates to each other, and so its relative position of significance among the whole. But the representation of this form is nothing other than our concept of happiness. So in judging that an object will positively contribute to the satisfaction of his total set of needs, and so will further the activity of his animal life as a whole, he is essentially judging it to fit with his concept of happiness. This concept thus underlies the relation between the object and the subject, making it the principle of judgment.<sup>33</sup> And when an object is judged in this way, as fitting with the subject according to the principle of happiness, then it is, at the same time, being judged to be *good*, or what I will call *agreeable in general*.<sup>34</sup> But I am still only judging it to be *good for* satisfying my overall needs, and so as mediately or conditionally good (rather than as immediately, objectively or absolutely good, as will be the case with the moral feelings). So these feelings are excluded from moral motivation. (A more detailed discussion of this point will be presented in the next chapter.) Finally, because general feelings of self-love involve the concept of happiness in this way, they can also be called *feelings of happiness*.

The involvement of happiness also explains the different way in which general feelings motivate. The motivational power of any feeling stems from the fit it represents. In

---

<sup>33</sup> “...the concept of happiness *everywhere* underlies the practical relation of *objects* to the faculty of desire...” (C2 5:25).

<sup>34</sup> I describe the predicate of these judgments in this way because Kant has three major categories of subjective predicates—the *agreeable/disagreeable*, the *beautiful/ugly*, and the *morally good/bad*. Objects of a judgment of happiness will definitely fall into the first category, and so will be a form of *agreeable*. But they are still very different from objects of a particular feeling, so in order to distinguish them, I will qualify them as *agreeable in general*.

the case of particular feelings, it was the expectation of *gratification* (i.e., the satisfaction of a particular need by the object). With general feelings, however, it is the expectation of *happiness* (i.e., an increase in the total gratification of needs by the object).<sup>35</sup> This follows because it is only through maximizing the possible satisfaction of needs that the total life of the animal is furthered. The ability to increase one's state of happiness (i.e., the degree to which one's sum total of needs have been satisfied (C2 5:124)) is thus the condition of an object being able to determine the faculty of desire. Because of this condition, I am said to act *for the sake of happiness*. To put this another way, I am acting from a representation of the value an object has in relation to my overall well-being as I conceive of it. When I represent the object in this way—*as agreeable in general* or *happiness-producing*—the concept of happiness is embedded within it. So when the representation of the object determines the faculty of desire, it specifically determines the *will*, because it does so in terms of the concept of happiness it contains as the condition for it motivating. The result is an *inclination to happiness*—an inclination that will be an expression of the harmony among needs brought about by the structuring effect of a concept of empirical practical reason.

But how exactly do judgments of happiness determine which objects are agreeable in general? This is essentially a question of how our concept of happiness is constructed. According to Kant, we start with an ideal—the total satisfaction of all our needs. We then fill out this ideal by incorporating more determinate parts established through experience. The process is thus one of building up a concept from experience, through the activity of

---

<sup>35</sup> "...a rational being's consciousness of the agreeableness of life uninterruptedly accompanying his whole existence is *happiness*..." (C2 5:25).

reason and the imagination, to complete the ideal.<sup>36</sup> For example, I learn from experience what my tolerance level is for coffee, for exercise, what amounts to appropriate hours of sleep, all by how I feel when I get too much or too little. I learn the physiological effects of caffeine on my bone health, blood pressure, etc., through science. And I even discover my physical skills or talents by trying out different sports, and seeing which ones I am more suited to and enjoy. Coordination of these then helps to generate my concept of health—a mid-level harmonizing concept. At the same time, I generate other mid-level concepts that connect with further needs or activities, such as talent development, leisure activities, financial stability, even temperament formation (moderating my passions). (These produce general maxims of action, or “little rules of prudence” akin to “little categorical imperatives”.) Coordination among these mid-level concepts then helps to form my overall concept of happiness. If I am someone who likes to travel, then I might have to develop certain talents over others in order to find the type of employment that would support it. Or if maintaining friendships is important to me, then I might have to adjust my career ambitions. Because of this imperfect, bottom-up process, our concept of happiness will always remain to some degree indeterminate and in constant flux as we come to discover, develop, and/or adjust our overall needs and goals. But it is also a concept that we build up in a way to *accurately* represent the totality of our subjective condition, i.e., the sum total satisfaction of our actual needs (present and future), given our actual talents, skills, goals and circumstances, whether we are aware of them or not. As such, it is a concept that we

---

<sup>36</sup> “In the idea of happiness, on the contrary, we have no concept of the whole, but rather we *only compose it out of parts*. And just for this reason we cannot direct our actions according to an idea of happiness, because such a whole cannot be thought by us” (LPDR 28:1057, see also G 4:418).

can get wrong.<sup>37</sup> So it is still an objective concept. (It is also a subjective concept because it represents the condition *of the subject*, and so is true only of him. Being subjective in this sense is what helps to make it a *practical* concept.) These two features thus explain how the concept of happiness can be *a rule* possessing normative force, but yet not a law that can govern absolutely (C2 5:36, and G 4:418). As an objective concept, it can tell us what we *ought* to do if we want to be happy. But because it is still indeterminate, it cannot state this claim absolutely.<sup>38</sup>

The principle of happiness thus has some degree of normativity, and so can act as a *regulative principle*. It has this function because it is a distinct principle of evaluation from the particular principle of self-love. As a second, more objective principle, however, it can come into conflict with, and so potentially override, judgments based on the particular principle of self-love. Consider the cases of Kate and Jake, both of whom are suffering from gout. Maintenance of their condition requires a certain type of diet—limited alcohol, red meat, and seafood consumption (what I described earlier as a rich meal). Now, while on a business trip to Chicago, Kate’s colleagues insist on eating at one of its famous steakhouses. If she just happens to not like meat, or has committed to being a vegetarian (for reasons other than her health), then she will be naturally inclined to one of the vegetarian options on

---

<sup>37</sup> This is why Kant states that “the greatest concern which he has for himself here would be... that he forms his concept [of happiness] falsely...” (NF #7199 19:272-3/462-3).

<sup>38</sup> “Happiness is the satisfaction of all of our inclinations (*extensive*, with regard to their manifoldness, as well as *intensive*, with regard to degree, and also *protensive*, with regard to duration). The practical law from the motive of **happiness** I call pragmatic (rule of prudence); ... [It] advises us what to do if we want to partake of happiness.... [And it] is grounded on empirical principles; for except by means of experience I can know neither which inclinations there are that would be satisfied nor what the natural causes are that could satisfy them” (C1 A806/B834).



the menu, an option that will also meet the dietary requirements of her condition. In such a case, there is no conflict between the judgments of her particular feeling and her feeling of happiness, because the meal she chooses will satisfy both her particular need for a tasty meal and her general need to maintain her health (and even her work relationships). She can then act for the sake of both gratification *and* happiness.

But what happens when the object of the two principles is not evaluated in the same way, causing the two principles to conflict? This is where the regulatory function of happiness becomes most evident. Jake, unlike Kate, has a particular love for a well-prepared steak. His selection of a nice Porterhouse will thus satisfy his particular need, and so will be judged to be *agreeable*. But as a significantly sized portion of red meat, it will negatively fit with his general need for health, and so will be judged to be *disagreeable in general*. (This is the situation in which the man in the *Groundwork* finds himself.) He now has two distinct and conflicting sources of motivation. Somehow these must be reconciled before he can act.

In general, there are three ways this conflict can be resolved. The first is for Jake to recognize the normative authority of the principle of happiness (as a concept of practical reason) over that of particular self-love, because the former aims to provide for the greatest possible satisfaction of the subject's overall needs (the ultimate end of his lower nature) (R 6:6n, and G 4:415-6). As a result of this recognition, he will be required to subordinate the particular principle of self-love to the principle of happiness *anytime* the two conflict. And when he does, he will then be able to resist or reject any suggestions contrary to his

happiness, such as an urge to order steak stemming from his particular feelings.<sup>39</sup> However, once the steak is rejected, the vegetarian dish then becomes the next best option in relation to his current particular need of hunger. Since this dish also fits with his general need for health, it will be the menu option he chooses. It is in this way that the principle of happiness functions as a regulatory principle, helping to bring our needs into harmony with each other. It thus follows that empirical practical reason governs our lower nature from within, as the evaluative principle in feelings of happiness.<sup>40</sup>

Now, if Jake's concept of happiness were fully determinate and accurate, then any failure to act according to the general rules it prescribes would be a form of practical irrationality. In such a case, Jake would be acting against what he *knows* to be his greater good. But as a concept that we build up from experience, it is never fully determinate and accurate in this way.<sup>41</sup> Its normative authority is weakened as a consequence. It cannot command in this state, but merely advise. This allows Jake a second option for reconciling

---

<sup>39</sup> That is, he will be able "to order his actions not in accordance with instinct but in accordance with concepts of happiness which he makes himself" and so avoid his chief concern, which is that he might "deviate from [his concept of happiness] through animal sensibility..." (NF #7199 19:272-3/462-3).

<sup>40</sup> "Everything is practical that is possible through freedom. But if the conditions for the exercise of our free choice are empirical, then in that case reason can have none but a regulative use, and can only serve to produce the unity of empirical laws, as, e.g., in the doctrine of prudence the unification of all ends that are given to us by our inclinations into the single end of **happiness** and the harmony of the means for attaining that end constitute the entire business of reason, which can therefore provide none but **pragmatic** laws of free conduct for reaching the ends recommended to us by the senses, and therefore can provide no pure laws that are determined completely *a priori*" (C1 A800/B828). "In fact, to the extent that practical reason is taken as dependent upon pathological conditions, that is, as merely regulating the inclinations by the sensible principle of happiness, this demand could not be made on speculative reason" (C2 5:120).

<sup>41</sup> "But it is a misfortune that the concept of happiness is such an indeterminate concept that, although every human being wishes to attain this, he can still never say determinately and consistently with himself what he really wishes and wills. The cause of this is that all the elements that belong to the concept of happiness are without exception empirical, that is, they must be borrowed from experience, and that nevertheless for the idea of happiness there is required an absolute whole, a maximum of well-being in my present condition and in every future condition" (G 4:418).

his two feelings. Because the normative authority is not absolute, he can question, and even override, the claims of his concept of happiness, and without crossing the line into irrationality. In fact, he must be able to do so as part of the process of self-discovery that enables him to build up and adjust his concept. More specifically, when a particular feeling comes into conflict with it, he must have the option to suspend the claim of happiness, and act on the particular feeling anyway. It is only in doing so that he could gain the type of experience needed as data to fill in, adjust or reinforce his concept. That is, Jake might try having the steak even though happiness warns against it. By doing so, he will be able to test just how much he needs to moderate his intake to prevent the painful consequences. Suspending the claims of happiness in this way can even lead to radical alterations in how we conceive of our lives progressing as a whole. But regardless of the degree, the goal for these changes is to make our representation of happiness more accurate and more determinate.

This brings us to the third option. Jake could suspend the claim of happiness “just this once” in order to act on the particular feeling. That is, more often than not, when he questions the claim of happiness, he is simply being tempted to make an exception to it through a process of rationalizing or self-deception. Particular feelings tend to encourage self-deception because the satisfaction they promise is both more immediate and determinate than that of happiness. And this can then give them a certain command over our attention that makes it possible to exploit the indeterminacy of happiness. For example, Jake’s desire for the steak might be so strong (and the expectation of pleasure so determinate) that he will be encouraged by it to adjust the expectation of future pain down from *probable*, to *merely*

*possible*. Or he might misrepresent the degree to which he must *limit* his intake of meat, saying “just tonight, because it is a special occasion—one time will probably not affect me that much.” It is this tendency that Kant takes particular note of in the *Groundwork*.

However, the precept of happiness is often so constituted that it greatly infringes upon some inclinations, and yet one can form no determinate and sure concept of the sum of satisfaction of all inclinations under the name of happiness. Hence it is not to be wondered at that a single inclination [inclination resulting from a particular feeling], determinate both as to what it promises and as to the time within which it can be satisfied, can often outweigh a fluctuating idea, and that a man—for example, one suffering from gout—can choose to enjoy what he likes and put up with what he can since, according to his calculations, on this occasion at least he has not sacrificed the enjoyment of the present moment to the perhaps groundless expectation of a happiness that is supposed to lie in health. (G 4:399)

The problem that concerns Kant, however, is *not* that a person might be tempted on occasion to act imprudently. It is that acting imprudently will allow the particular feelings to exaggerate their own value and so eventually turn into passions. In other words, by indulging these particular feelings, their influence will strengthen. And this will, in turn, lead him to adopt a corrupted concept of happiness through the process outlined in the second option. Kant expresses this concern directly in one of his notes. “Inclinations, united through reason, agree with happiness, i.e., with well-being from the enduring satisfaction of all our inclinations. Single inclinations, if they hinder attention to the satisfaction of the remaining ones, contradict happiness. Passions thus naturally contradict not just morality but also happiness” (NF #6610 19:107/422). When this corruption occurs, it will lead to sadness, and eventually to vice. It is in order to help guard against this that Kant raises the notion of happiness as an indirect duty.

#### **IV. Sadness and the Passions**

In the last section, we saw that there is a collection of elements involved in Kant's theory of happiness. Actually being in a state of happiness, for Kant, is the degree to which one's overall needs have been satisfied. The concept of happiness is a representation of that totality of needs, and the relative importance each has to the others. When this concept is fully determinate and accurate, it will reliably guide us to the maximal satisfaction of needs possible for us given our finite condition and limited resources. When this concept becomes the principle of a judgment of feeling, it then sets its object as an end of the will. This results in an inclination to happiness, which then motivates us to make its object actual. When successful, the need will be satisfied, resulting in an increased state of happiness. And since we generate our own concept of happiness, one that is tailored to our particular subjective condition (our individual set of wants, needs, skills, talents, goals, interests, and circumstances, etc.), we essentially become, through this concept, the authors of our own happiness. That is, we begin to set our own ends apart from those that nature has set for us. So it is in structuring our needs through the feelings of happiness that we essentially increase our freedom, because this is how our animality is transformed into humanity. This is why Kant states that "As a freely acting being, indeed in accordance with this independence and self-rule, he will thus have as his foremost object that his desires agree with one another and with his concept of happiness, and not with instincts; and the conduct befitting the freedom of a rational being consists in this form. First, his action will have to be arranged in accordance with the universal end of humanity in his own person and thus in accordance with concepts and not instincts..." (NF #7199 19:273/463).

Now, if this is happiness, then what is *sadness*? Kant gives a brief description in the *Lectures on Anthropology*, indicating that it is the contrary of this full picture of happiness.

Pain is mere sensation. Sadness [is] when I consider myself unfortunate. It is a judgment on the worth of my entire condition. But pain [is] merely a partial sensation. –A man of principles can never be unfortunate. He can surely feel pain, but not sadness. This latter is a new pain that arises from one's feeling that his pain outweighs his entire enjoyment. (LA 25:1325)

Sadness is related to pain in the way that happiness is related to gratification. When a need is satisfied, the subject feels the sensation of gratification as a result. When a need is thwarted, the subject feels pain instead. The more our set of total needs are satisfied, the greater our happiness (sum total of gratification). And now we see that the more our set of total needs remain unsatisfied, the greater our feeling of sadness (sum total of pain). So just as happiness comes about from a harmony of needs, sadness likely comes about from a disharmony of needs, one that in turn decreases the chance of their total satisfaction. But if reason is responsible for the harmony of happiness, then what is responsible for the disharmony of sadness? In a word, the *passions*.

For Kant, the passions are the subset of feeling that is inimical to our own freedom, and so to the control of reason. Because of this, he perceives them as a serious threat to both our happiness and to morality. In the *Anthropology* he describes passion as an “inclination that prevents reason from comparing it with the sum of all inclinations in respect to a certain choice... It is also easy to see that they do the greatest damage to freedom ... [it] is an enchantment that also refuses recuperation” (A 7:265). In the *Religion*, they are inclinations that “exclude mastery over oneself” (R 6:29\*). And in the *Lectures on Anthropology*, he states “If an inclination for a thing has become so strong and habitual that it suppresses all

other inclinations, then it is called a passion. With passion, one is not able to compare the inclination with the sum of all other ones. Passion is blind in that it does not want to adopt the judgment and conclusion of the understanding” (LA 25:1339-40). In essence, passions are the result of particular feelings that have become so strong that they now completely resist comparison with other needs, and so with the concept of happiness (and morality). And when they do, they undermine the very freedom and self-determination that happiness affords us. In other words, when we allow our passions take over, we are, in essence, putting ourselves in chains.

But how do particular feelings get to this point? In large part, it is their nature as *particular* feelings that allow them, when left unchecked, to develop in this way. Surprisingly, Kant takes up this issue in, of all places, the *Lectures on Logic*. There he gives an interesting description of inclination (need) as the ground of a sort of provisional judgment (particular feeling) requiring reflection before being accepted as true.

*Inclination* is also a cause and ground of some judgments, furthermore, and in this case, to be sure, one judges and infers that something is good, acceptable, and perfect because it excites us, because it stimulates us and suits our taste.

Still more, *inclination* occasions us always to undertake examinations and investigations only from one side, and of course only from the side where we wish that it were so and not otherwise, and thus it occasions us to leave the other side, which might perhaps provide us with grounds for the opposite, completely uninvestigated. We even seek to find grounds with whose help we can refute the opposite of what we wish. Here, likewise, the laws of the understanding and of reason are, as it were, absent and not at home; there is simply no reflection.... (LL 167/132)

A particular feeling is thus a provisional judgment—a judgment that an object is good in relation to a single need. But before we can accept this judgment, we must compare it to our concept of happiness (or morality) to make sure that the opposite might not actually be true.

We need to do this for two reasons. First, because these are *provisional* judgments, and so are prone to error, inflation of their claims, etc. And second, because particular feelings suggest something that we want to be true, we will tend to find them very convincing. (This is why Kant often describes them as being *seductive*). But because they are so convincing, we often fail to reflect on them. When we do, they become what he calls a *prejudice* or a general rule for judging a certain way (but still without reflection to make sure it is correct) (LL 143-5, LL 161-7, LL 737-40 and LL 58-76).<sup>42</sup>

Particular feelings thus start out with a tendency to prevent comparison with the concept of happiness. And even when we try to engage in reflection, they can distort our attempts by emphasizing the positives that would count in its favor and de-emphasizing the negatives that would count against it. I described this tendency in the previous section as engaging in a process of self-deception and rationalization. Because Jake was susceptible to its persuasion, he was tempted to make an exception to the claim of happiness and order the steak anyway. Now, an exception here and there is not in itself a problem. But Kant's worry is that this tendency, when coupled with the process for adjusting our concept of happiness discussed in option two, will lead to the concept's total corruption. That is, when we indulge these needs, they will intensify, which then leads to an inflation of their value or significance relative to the total set of needs. Jake will then adjust his concept to reflect its new status. When he does, they will again strengthen, encouraging Jake to make additional exceptions to the rule. If left unchecked, this process will continue until it becomes the

---

<sup>42</sup> This is why Kant is concerned about acting from *habit*. He believes that the more habitual an action becomes, the less we reflect on the judgment motivating it before we do, and so the more error prone our actions will become (LA 25:1334).



totality of Jake's concept of happiness. He then becomes incapable of correction because the standard by which he would do this has itself become corrupted.<sup>43</sup> Jake is then in the grips of a full-blown passion. Kant describes this tendency of the passions in relation to ambition in a passage in the *Anthropology*.

Passions are cancerous sores for pure practical reason, and for the most part they are incurable because the sick person does not want to be cured and flees from the dominion of principles, by which alone a cure could occur. In this sensibly practical too, reason goes from the general to the particular according to the principle: not to please one inclination by placing all the rest in the shade or in a dark corner, but rather to see to it that it can exist together with the totality of *all* inclinations. –The *ambition* of a human being may always be an inclination whose direction is approved by reason; but the ambitious person nevertheless also wants to be loved by others; he needs pleasant social intercourse with others, the maintenance of his financial position, and the like. However, if he is a *passionately* ambitious person, then he is blind to these ends, though his inclinations still summon him to them, and he overlooks completely the risk he is running that he will be hated by others, or avoided in social intercourse, or impoverished through his expenditures. It is folly (making *part* of one's end the *whole*) which directly contradicts the formal principle of reason itself. (A 7:266).

This is why Kant uses such hostile language when it comes to the passions. In general, they refuse to have their value or significance determined by their relative position with respect to the total set of needs. Instead, they insist on asserting their own value *independently*. When they do, they are basically seeking to disrupt the harmony reason attempts to impose in order to inflate their own status. With the elevated status, they then demand satisfaction at the sacrifice of other needs, bringing about even more disharmony. When a passion becomes strong enough, it can assert itself as constituting the entirety of the subject's concept of happiness.

---

<sup>43</sup> This corruption happens on two levels—first with the principle of happiness, and then (in a different way) with the principle of morality. I discuss the second way in the next chapter.

The result of this corruption of the concept of happiness is thus sadness. Why? Because happiness does not result when the subject satisfies the needs he thinks he has, but only when he satisfies the needs he actually does have. This is why it is so important to have an accurate concept of happiness and to reflect on the suggestions of particular feelings before acting. With Jake's corrupted concept, he now aims to satisfy a limited set of needs, even when he is acting from a feeling of happiness. In other words, his judgments of happiness will be faulty, over-valuing certain objects in relation to his needs, while under-valuing (or not even valuing at all) others. This then leaves a greater proportion of his needs unsatisfied, resulting in a state of sadness.

It also leaves him with an impaired ability to set his own ends, and so less free. When a passion takes over, he is no longer acting from his conception of happiness, but has essentially turned the reins of control back over to nature. His needs now determine him, rather than the other way around. This is why Kant describes the state of sadness as a kind of insanity, because it involves the corruption of a concept of reason, which then leads to a renunciation of his own freedom. "Only in some cases does he have no power of free choice, e.g., in the most tender childhood, or when he is insane, and in deep sadness, which is however also a kind of insanity" (LM 28:255).

We now see why Kant emphasizes that we need to be the authors of our own happiness, because it is in doing so that we finally become free. And it is neglecting our happiness that we can put ourselves back into chains.

## Chapter 3

### Respect for the Moral Law: the Emotional Side of Reason

There is little doubt that Kant saw respect as a central piece of his moral theory. It might even be considered the linchpin, because by means of it and it alone reason becomes practical. But as Kant describes it, respect has a duality of nature that seems to embody a contradiction; namely, respect is both a moral motive *and* a feeling.<sup>44</sup> And given the basic metaphysical and anthropological framework of his moral theory, where there exists a deep division between the moral/rational and natural orders, it seems that these two aspects of respect land on opposite sides of the metaphysical divide. As a moral motive, respect must be purely in the rational order, unadulterated by sensibility. But in order to be a feeling, that is, to be *felt*, sensibility must be involved. So respect, as a feeling, cannot be a moral motive.

Given how deeply incompatible these two characteristics appear to be, it is not surprising that Kant's account of respect has been considered an extremely problematic element of his theory. In response, most sympathetic commentators have suggested that Kant never really intended for respect to be understood as both a moral motive *and* a feeling. Thus, most attempts at resolving this issue have focused on eliminating this duality. Rather

---

<sup>44</sup> References to respect as a moral motive and as a feeling are scattered throughout Kant's major moral works. Consider as examples: "Respect for the moral law is therefore the sole and also the undoubted moral incentive" (C2 5:78), and "It could be objected that I only seek refuge, behind the word *respect*, in an obscure feeling, instead of distinctly resolving the question by means of a concept of reason. But though respect is a feeling, it is not one *received* by means of influence; it is, instead, a feeling *self-wrought* by means of a rational concept and therefore specifically different from all feelings of the first kind, which can be reduced to inclination or fear" (G 4:401n).

than see respect as a single mental state with dual qualities, they instead bifurcate it into distinct, albeit closely connected, states involved in the determination of the will, only one of which is *properly* called respect. On the one side, we have the consciousness of the moral law and/or the recognition of its authority (a purely rational, cognitive state). On the other, we have the effect of this cognitive state on our sensibility in the form of a painful thwarting of inclination (a natural/empirical, affective state). The two dominant views diverge in regards to which of the two states they identify as respect. Andrews Reath is one who argues that respect is consciousness of the moral law, and so the moral motive, whereas Paul Guyer identifies respect with the pathological aftermath, and so views it as a feeling.<sup>45</sup>

Though these are both interesting accounts, I think they are ultimately unsuccessful because the basic strategy they both pursue for resolving the problem fails to capture what is unique about respect, and therefore fails to understand *how* respect motivates. Specifically, I believe that Kant thinks respect can function as the moral motive *because* it is a feeling, not despite it. So rather than eliminate the unique dual character of respect, I suggest we preserve it, and instead explore Kant's view of emotion giving rise to it. For Kant to make such a bold claim about respect, he certainly must have had a fairly substantial theory operating in the background (whether or not it ever manifested itself explicitly in his main works). And I believe he did—it will just take some philosophical excavation to uncover it and piece it together. Moreover, I believe his various discussions of respect are places

---

<sup>45</sup> I believe Reath and Guyer represent clear examples of these two views, but they are not alone. Most accounts of respect fall into one or the other. For more on these specific accounts, see Andrews Reath, "Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination," *Kant-Studien* 80 (1989): 284-302, and Paul Guyer, *Kant on Freedom, Law, and Happiness* (New York: Cambridge University Press, 2000).

where he was working out one of the more difficult elements of his theory—specifically, where he was attempting to merge the motivational aspect of emotion with the objectivity of pure reason.

In Chapter One, I argued that Kant held a cognitive theory of emotion, i.e., a theory in which emotion is at least partially constituted by reason, rather than the non-cognitive or mechanistic theory often attributed to him. This change alone was a step forward because it allowed reason at least some ability to shape or influence non-moral motivation. But it is not yet enough to solve the problem at hand. In order to account for *respect*, emotion will have to take on a *purely* cognitive/rational form, a form absent any involvement from sensibility. How does Kant achieve this end? By characterizing emotion primarily in terms of its *function*, i.e., in terms of its characteristic role in the activity of the mind. More specifically, emotions are *action-initiating evaluative judgments*, a position revealing a deep Stoic influence. The forms, objects, principles and characteristics of these judgments will vary across a wide spectrum, from the purely non-rational/non-cognitive *analogues* of judgment on the instinctual end,<sup>46</sup> to the purely rational/cognitive judgments on the moral end, and passing through the usual evaluative judgments of ordinary emotional responses in between. With this general framework in place, Kant can then account for the emotions of any living, animate being, from non-rational animals all the way up to a purely rational, non-

---

<sup>46</sup> This type of judgment accounts primarily for emotion in animals. As non-rational creatures, they only have an analogue to reason and its judgments, and consequently only produce a “connection of representations according to the laws of sensibility, from which the same effects follow as from a connection according to concepts” (LM 28:276, see also LM 28:690). Most human emotion, in contrast, will involve reason to some degree. (Even though our concept of *emotion* is slightly less expansive than that of *feeling* for Kant, I will use the two terms interchangeably. I chose to do this because ‘emotion’ conveys the same general sense in the relevant passages but without the questionable emphasis on sensation, a concern about ‘feeling’ (*Gefühl*) Kant shared as well (LPDR 28:1059).)

embodied god. And since the key functional feature of emotions is that they are *action-initiating* judgments, emotions emerge as the centerpiece of Kant's motivational account. Reason becomes practical by becoming emotional.<sup>47</sup> It follows, then, that there is in fact *no* motivation in Kant without emotion—moral or otherwise.

## I. The Faculty of Feeling of Pleasure and Displeasure

We saw in the previous two chapters that the account of emotion routinely attributed to Kant is non-cognitivism. With this theory, feelings are primarily defined in terms of their qualitative or affective characteristics, i.e., what they *feel* like. I argued instead that Kant understood feelings of pleasure to be action-initiating evaluative judgments. In doing so, he essentially defines emotion in terms of its function—i.e., its characteristic role in the activity of the mind. When feelings are understood in this way, pleasant sensations and evaluative judgments can both fall under this category because they have the same effect in motivation—they initiate action, or more specifically, they *promote the life or activity of the subject*. That Kant would describe pleasure in terms of its function, rather than in terms of its affective character or visceral feel, should be no surprise. He strongly believed that there could be various forms of life, with radically different types of sensibility. Animals, for

---

<sup>47</sup> This claim reflects the Stoic thesis in which the rational and emotive faculties are identified. I believe that this thesis, along with the structure of purely rational emotions (*eupatheiai*) it makes possible, are the two central points of Stoic influence on Kant's own theory of emotion. Kant expresses his enthusiastic approval of the Stoic account when he states that they "sowed the seed for the most sublime sentiments (*Gesinnungen*) that ever existed" (LL 9:30/542). However, though Kant adopts the Stoic constitutive thesis that emotions are evaluative judgments (and not the sensations sometimes associated with them), he will reject the normative thesis that all ordinary emotions are false judgments and so should be suppressed. On the various Stoic theses, see Margaret Graver's *Stoicism and Emotion* (Chicago: University of Chicago Press, 2007). Kant's own theory also echoes Augustine's advancement of the Stoic line, in which he argues that emotions are acts (judgments) of the will. See the *De civitate Dei*, XIV.6.

example, have no inner sense. Angels, as spiritual beings, have no outer sense. Saturnians and Mercurians are likely to have versions of the inner and outer senses vastly different from ours. God, unlike a finite creature, is a purely active being and so has no senses at all (LM 28:275-8, A 7:141fn24, LM 28:211 and LPDR 28:1051). This variety of sensibility (or lack thereof) poses a challenge, because the faculty of feeling is a faculty Kant believed to be shared by *all* minds, even God's, because it plays an integral role in the activity of a mind as such (R 6:73, LPDR 28:1056 and LPDR 28:1059-61). So he must be able to characterize the faculty in a way that is independent of the particular form of life a being might have, and he does this by describing pleasures in terms of their functional role in motivation. If, instead, Kant had described pleasures in terms of a felt, experiential quality, then God would have the faculty of feeling without being able to feel, among other problems.<sup>48</sup> So just as this might suggest, understanding emotion in terms of its function is what will enable Kant to give an account of purely rational emotions suitable for moral motivation.

At this point, we know feelings have two basic features. First, they are defined in terms of their function. That is, *any* representation capable of motivating action (i.e., determining the will) is, by Kant's definition, a feeling. Second, what enables these representations to motivate is the underlying structure they all share. Feelings are judgments with a particular evaluative form.

---

<sup>48</sup> This is not to say that emotions could not have a connection to the body, or depend in some way on a particular being's sensible constitution. To the contrary, certain feelings on Kant's view will require an affective component in order to meet the functional criterion. Instead, it is simply a denial that this affective component is part of the *essence* of feeling or emotion in general. Kant's view is that *any* motivationally efficacious evaluative state that has the good as its object will be a feeling, whether or not it is felt. By not requiring sensation, Kant is thus able to attribute emotion to purely rational/spiritual beings, and so also to our own rational nature.

Kant describes both of these features of feeling in an early passage of the second *Critique*, where he states

...the determining ground of choice is then the representation of an object and that relation of the representation to the subject by which the faculty of desire is determined to realize the object. Such a relation to the subject, however, is called *pleasure* in the reality of an object (C2 5:21).<sup>49</sup>

In other words, pleasure is a representation of an object, not as it is in itself, but as it *relates to the subject*. Or more specifically, it represents how the object relates to the subject's faculty of desire—the source of his activity.<sup>50</sup> Without this connection to the subject's source of activity, an object cannot move us to act. It is in this regard, where pleasure is a type of judgment representing this relationship, that Kant says the faculty of desire is determined by a pleasure in an object. Motivating action (determining the will) is thus the function of pleasure.

But perhaps the most striking discussion of this functional characterization of pleasure can be found in the *Lectures on Metaphysics*, where Kant is discussing the moral feelings—feelings that he admits would stretch the use of the term *feeling* beyond its ordinary limits.

---

<sup>49</sup> In this passage, Kant is discussing pleasure in connection to the material principle of self-love. Despite the context in which it is presented, this definition holds more broadly. See, for example, LM 29:894, where Kant says that “Pleasure is the representation of the agreement of an object with the productive power of the soul [the faculty of desire], and displeasure the opposite.” Or NF #1021 15:457/408, where he states that a “representation must...have a relation to the subject of determining it to action. This relation is a pleasure....” (See also C2 5:9n, MM 6:212, C3 5:209, LPDR 28:1060 and NF #715 15:317/495.)

<sup>50</sup> Kant refers to the subject of these judgments in a number of ways, including the subject's *faculty of desire*, *life*, *productive power of soul*, and *freedom*. These are all references to the source of causality within the subject—what makes the subject an *active* or *living* being (LM 28:247, LM 29:891 and C2 5:9n).



I am supposed to have a feeling of that which is not an object of feeling, but rather which I cognize objectively through the understanding. Thus there is always a contradiction hidden in here. For if we are supposed to do the good through a feeling, then we do it because it is agreeable. But this cannot be, for the good cannot at all affect our senses. *But we call the pleasure in the good a feeling because we cannot otherwise express the subjective driving power of objective practical necessitation* (LM 28:258, my italics).

In this passage, Kant is saying that even though the good is not something I cognize through the senses (and hence involves no sensation or felt quality), it is still a feeling because this cognition has a *subjective driving power*, i.e., it motivates.<sup>51</sup>

Consider, for example, watching a storm move in over Lake Michigan. That stormy weather can produce large waves and fresh, cool air does not in itself motivate me to head to the lakefront for a walk. But when I hold a particular fascination with the changing moods of the lake, and recognize that I have been inside all day writing and could use some fresh air, a brewing storm *does* motivate me to head outside. The stormy air and the crashing waves will help to clear my mind and stimulate new ideas. Kant describes this relationship between the storm and my subjective condition as a sort of “fit” that promotes my life or activity—in this case, my philosophical writing. The representation of this fit, in the form of *fascination* and *excitement* at the wonders of nature, is the pleasure that determines my will. Furthermore, when I judge watching a storm in this way, as positively fitting with my needs and activities, then I am at the same time judging it to be *good* (i.e., to be *pleasing*). These pleasures, then, are a type of evaluative judgment that involves the subjective predicates of

---

<sup>51</sup> The functional character described here as a subjective driving power is expressed elsewhere as the *promotion or furtherance of life*, and as a *ground of an impulse to activity* (where the impulse itself is a desire) (LM 28:247, G 4:422, LL 24:45/31 and NF #5448 18:185/415). All of these descriptions are attempts to capture pleasure’s power to motivate.

*good/bad*—or more specifically, of *agreeable/disagreeable* and (moral) *good/evil* (where the use of *evil* distinguishes the moral good from the generic good) (LM 28:245).<sup>52</sup>

The same holds for displeasure. If instead I were out sailing when the storm approached, I would likely feel *fear* or *anxiety* rather than *excitement*. The strong winds and high waves would increase the chance of capsizing, putting my life at risk—i.e., they would negatively fit with my needs and activities. As a result, I would judge the storm to be *bad* (i.e., *displeasing*), and immediately head to shore to avoid it. Any object that neither pleases nor displeases me in this way (neither promotes nor hinders my activity) leaves me motivationally disengaged. I am *indifferent* to it because it would have *no* connection to my will at all, the source of my activity (LM 28:253). So unless a representation, of whatever kind (from physical objects to moral concepts), is connected to the faculty of desire in this way, it cannot motivate. But any such connection is a feeling. It thus follows that *all* motivation must go through the faculty of feeling, and so must take the form of a pleasure—even the moral motive of respect. And insofar as feelings are *judgments that determine the will*, I believe Kant understands them to be practical cognitions, which he defines as “cognition[s] having to do only with the determining grounds of the will” (C2 5:20, see also C1 Bix-x and LL 24:58/42).

We now have a general sketch of Kant’s functional account of feeling. But where it is perhaps most evident is in the role he sees the faculty of feeling playing in the overall structure of the mind. This role has been obscured because most of us are accustomed to

---

<sup>52</sup> The *beautiful/ugly* are also subjective predicates, but since they are not directly connected to motivation and action (i.e., are not practical pleasures), I will not discuss them here.

thinking in terms of a bipartite structure, where the basic functions of the mind are divided between its cognitive and desiderative capacities. However, Kant has a tripartite theory, where the faculty of feeling is considered to have equal standing with cognition and desire. As a major faculty, feeling will have its own characteristic function. And this function depends in large part on Kant's understanding of the activities and limitations of the other two faculties, and the connection required between them.

In its most general characterization, the faculty of cognition is, for Kant, the faculty responsible for objective or theoretical cognition. This faculty allows us to distinguish objects by means of theoretical predicates, such as shapes and colors, which are properly attributed to the object. The predicates of *good* and *bad*, however, must be excluded from this faculty because, as evaluative predicates, they are *not* properly attributed to the object. If the object itself were pleasing/good, then the subject's agency would be compromised, because *any* time the object were cognized, the subject would be motivated to pursue it. This response to the object, however, would not really be an action, but instead a mere mechanical *reaction* to the perception of agreeableness, like the motion of sunflowers in relation to sunlight. So in order for there to be genuine action, the subject's response to the object must be able to take into account his own particular needs and circumstances, because it is only by doing so that he can act with a conception of his own individual happiness or well-being in mind. But, when his subjective condition is taken into account, the predicate *good* cannot then be attributed to the object itself, but to the object only when considered in relation to the subject. Hence, it cannot be a theoretical predicate (LM 28:246, LM 29:877-8, NF #823 15:367/506 and MM 6:211-13). (We will see, however, that the objective

pleasures of reason will prove to be an interesting exception to this.) As a result, the faculty of cognition is restricted to the theoretical judgment that a storm produces cool air and high waves. Any further judgment that it is good (i.e., satisfies a need) lies outside of this faculty's domain.

The faculty of desire, in contrast, is active over the domain of the good and bad, whereby we pursue the former and avoid the latter. The principle of this faculty is the familiar "I desire nothing but what pleases [the good], and avoid nothing but what displeases [the bad]". Of special note here, however, is what Kant immediately goes on to say, "But representations cannot be the cause of an object where we have no pleasure or displeasure in it. This is therefore the subjective condition by which alone a representation can become the cause of an object" (LM 29:894, see also LM 29:899-900). In other words, I cannot simply desire to head outside because I represent stormy weather. Rather, it is only after the cool air and high surf have been judged to be good (pleasing) that I can have such a desire. It follows that the function of the faculty of desire is not to judge an object to be good and thereby pursue it, but to desire/pursue it because *it has already been judged to be good*. (This distinction between judgments of the good and the faculty of desire is important, because without it, Kant would be unable to account for non-practical pleasures—pleasures that can be understood to promote certain mental activities such as aesthetic contemplation and the assumption of the postulates of pure practical reason (i.e., faith) (C2 5:119-120, LM 29:877-8, MM 6:211-14 and C2 5:142-6).) As a result, the activity of the faculty of desire is the actual movement toward the object to be brought about. Something else must initiate that movement.

As Kant describes cognition and desire, there is no overlap between them. The faculty of cognition can tell us that a storm produces cool air and high surf, but not that this weather is good for a stimulating walk. However, I can only desire the good. So without something connecting these two faculties, there is no way that I can come to desire a walk on the lakefront. I am left motivationally inert. Kant makes this point explicitly when he briefly considers this scenario. “If we take away the faculty of pleasure and displeasure from all rational beings, and enlarge their faculty of cognition however much, then *they would cognize all objects without being moved by them*; everything would be the same to them, for they would lack the faculty for being affected by objects (my italics)” (LM 28:246, see also LPDR 28:1065-6). By the phrase “to be moved by an object”, Kant means to be moved to activity (i.e., to be motivated). He is thus saying that the faculty of cognition alone cannot move the subject because, as noted earlier, theoretical cognition only registers the properties of the object. This leaves the faculty of cognition disconnected from the faculty of desire. But, as Kant envisions the mind, this connection is essential for motivation to be possible at all. The faculty of feeling fills this gap by being the faculty responsible for practical cognition—i.e., the faculty that judges the object of a representation of cognition to be good in relation to a subject, and so brings the representation under the active scope of the faculty of desire, our source of causality for making the object of that representation actual (LM 29:890). As the faculty responsible for practical cognition, we can now see why Kant elevates feeling to one of the three major faculties of the mind. It is the faculty central to Kant’s entire account of motivation, for moral and non-moral action alike.

## II. The Lower and Higher Faculties of Feeling

One of Kant's overarching goals in developing his moral theory is to find a non-empirical source of motivation. But where there is motivation, feeling must be involved. So corresponding to the two types of motivation, there must be at least two types of pleasure, one empirical and one non-empirical (LM 29:1024 and C3 5:205-6). This distinction in pleasure is reflected in the division Kant makes between the lower and higher sub-faculties of feeling. He describes this division immediately after making similar divisions in the faculties of cognition and desire.

Likewise the *faculty of pleasure and displeasure* is also a *higher* or *lower* faculty. The *lower* faculty of pleasure and displeasure is a power to find satisfaction or dissatisfaction in the objects which affect us. The *higher* faculty of pleasure and displeasure is a power to sense a pleasure and displeasure in ourselves, independently of objects. All lower faculties constitute *sensibility* and all higher faculties constitute *intellectuality*...--But *intellectuality* is a faculty of representation, of desires, or of the feeling of pleasure and displeasure, so far as one is wholly independent of objects (LM 28:228-9, see also LM 29:877, LM 28:252, A 7:141fn24, A 7:159fn53 and C3 20:245).

As described in this passage, the higher sub-faculties constitute intellectuality (i.e., self-activity or spontaneity), and the lower constitute sensibility. The corresponding sub-faculties of cognition are the understanding and sensibility; those of desire, will and choice. However, despite the familiarity of these divisions in cognition and desire, the parallel division in the faculty of feeling has been largely missed. (One reason for this might be that Kant never assigns specific names to the sub-faculties, referring to them instead by the objects and/or products of each—their objects as the good/evil and the agreeable/disagreeable, and their products as moral/intellectual feelings and

pathological/sensible feelings, respectively). But I believe this division is the key to understanding Kant's account of moral motivation. As a form of sensibility, we know that the lower faculty of feeling will be empirical. And as a form of intellectuality, the higher faculty of feeling will be non-empirical or rational. If this distinction holds, then the higher faculty will be well-suited to explain the feeling of respect as a moral motive. But first, Kant must give an account of these two very different types of feeling.

It is at this point that Kant's turn to a functional/judgment account of feeling reveals itself to be one of the crucial moves in developing his theory of motivation. These two radically different types of pleasure are possible because of the underlying functional structure they share. To be a feeling of any type, an object must be judged to fit with a subject in a way that furthers the subject's life or activity. If the object furthers life, then it (as the object of the pleasure) determines the will. Within this basic form, certain variations are possible. The object of pleasure can be anything from fresh air to a purely rational/moral concept such as truthfulness. The principle can be either the empirical principle of self-love, or the *a priori* principle of morality. Finally, the subject can be one of two principles of life in a finite, rational being—its animal (physical) life or its spiritual (rational) life (corresponding to the lower and higher faculties of desire) (LM 28:286, LM 28:248).

Animal life involves natural causality, or causality in relation to our physical nature.

Spiritual life, in contrast, involves freedom, or causality in relation to our rational nature.

With this general outline in hand, we can now turn to the individual sub-faculties, beginning with the lower feelings. Though this faculty has a rather complicated and interesting structure, one that I will not be able to do full justice to here, a quick sketch will

nevertheless provide an intuitive framework against which to contrast Kant's (less intuitive) account of the higher faculty.

### ***The Lower Faculty of Feeling***

The fascination and excitement felt in watching a storm discussed above is an example of a lower feeling. With these feelings, there is a sensible need, stemming from my animal nature, that drives the judgment—e.g., to clear my mind. Certain objects are considered in relation to these needs, some of which are judged to fit. When there is a positive fit, the principle of self-love is determined to hold between the object and the subject. And since a *sensible* need is the basis of this judgment, the resulting pleasure will determine *choice* (the *lower* faculty of desire) to make the object actual. Finally, when an object is determined to fit in this way, it is judged to be *agreeable*, the particular predicate Kant assigns to judgments of the lower feelings (along with *disagreeable*) (LM 28:248).

The basic structure of the lower pleasures also determines *how* they motivate. That is, their motivational power stems from the fit they represent. I judge that the approaching storm fits with my needs because the fresh air and mental stimulation it provides *will clear my mind* and so *help promote my writing*. Implicit in this representation of fit, then, is the *expectation of a need being satisfied* by the storm. It is this expectation that ultimately motivates me to head outside. Now, Kant distinguishes the actual satisfaction of my need, i.e., the mental clarity I come to feel, as a second type of pleasure—what he specifically refers to as sensations of *gratification/enjoyment* and *pain*. In the case of the lower faculty, these two types of pleasure (what I will call judgments and sensations of pleasures,



respectively) must work together to motivate. That is, the lower judgments of pleasure can motivate only because of the expectation of gratification (satisfaction of a need) they represent. This follows because it is only through the gratification of a need that our animal life is furthered. The ability to gratify is thus the *condition* of the stormy weather being able to determine the will.<sup>53</sup> Because of this condition, I am said to watch the storm *for the sake of* the expected gratification. So when I judge the storm as agreeable, I am judging it to be *good for* satisfying my needs. However, in doing so, the representation of the expected gratification always mediates between the subject and the object. Hence, the object is judged as *mediately good*, rather than *good in itself* (LM 28:252, LM 29:891, C2 5:59, C2 5:62 and C3 5:207). It is for this reason that Kant calls all lower pleasures forms of self-love—because they are all ultimately directed towards satisfying one’s own needs.

Given the description above, it might appear that the possible objects of the lower judgments of feeling would be restricted to physical objects or states of affairs. But this isn’t quite the case. The object can still be a representation of reason, even of morality—but with one caveat. The only way that such a representation can be judged to fit with the subject’s animal nature is by bringing a state of affairs about in the world to match the concept. This follows because the need governing the fit in the lower feelings between the concept of morality (the object) and the subject (his animal nature) requires sensible

---

<sup>53</sup> Clearly, stormy weather is not agreeable to everyone. Kant notes this by pointing out that the expected gratification that serves as the basis for my judging such weather to be agreeable is dependent on the privately valid grounds of my own senses (LM 28:248, LM 29:892 and NF #1512 15:836/525-6). I.e., because I happen to have a certain sensible constitution (having been raised in tornado alley), stormy weather has a positive effect on me. So this judgment holds *only for me*—i.e., it is *merely* subjective (C3 5:212 and NF #1850 16:137/536). For these judgments to be objective, in the sense of holding universally and necessarily for all beings (and so to be suitable for moral motivation), a different ground of validity will be required.

gratification. So even in the case of a moral concept, the resulting pleasure is an attitude directed not internally to the form of willing itself, but externally to the state of affairs that will result from the action. In other words, lower judgments of feeling can only provide the *matter* of the will. Because of this, I distinguish these feelings as *material* pleasures (as opposed to what will be the *formal* pleasures of the higher faculty). Material pleasures, then, are the specific pleasures that motivate in terms of expected gratification, whether or not they involve an intellectual representation (such as a moral concept).

Kant's case of the naturally sympathetic man provides an example here (G 4:398 and C2 5:34). The object of his pleasure might be the rational concept of *beneficence*. But for this concept to be an object of a lower pleasure, it must promise gratification in order for it to motivate. That is, the action represented by the concept must be judged to fit with the man's psychological need to spread joy around him, a need which is gratified only when he succeeds in making others happy.<sup>54</sup> If there were no such need, then the representation of the action would fail to motivate. His goal then is not to will well (by willing beneficence), but to bring about a certain state of affairs in the world by which his psychological need will be satisfied. In this situation, the moral concept would essentially become a means to this gratification, and therefore would be judged as *mediately good*, rather than *good in itself*.

This is the model of moral motivation (what I call the Epicurean model) that Kant is at pains

---

<sup>54</sup> Some have thought that Kant's description of the sympathetic man is problematic because it actually portrays him as selfish rather than altruistic. But to act from a psychological need does not necessarily entail selfishness. As I believe Kant sees it, the man is altruistic because of the sort of needs that he has—he has a need to help others, and he takes immediate satisfaction in doing so when he can. This is an ordinary description of natural altruism—someone who just enjoys helping others. A selfish person, in contrast, may have a need to help others, but does not *immediately* enjoy doing so. Instead, she helps others only as a means to some further end, such as a good reputation, and it is only this further end that she finds gratifying.

to reject in the second *Critique* (C2 5:23-26 and C3 5:208-9). The adherents to this model assume that the rational source of the motivating concept is enough to establish its moral status. But as Kant shows, if a judgment of pleasure, even in a moral concept, motivates only in terms of the gratification presupposed, then it is a form of empirical, material motivation unsuitable for genuine moral action.<sup>55</sup>

Kant's condition for moral motivation is thus clear—no feeling that presupposes gratification (i.e., no *material* feeling), regardless of the source of its representation, can be a moral motive. So the motivational force of the higher feelings will have to be based on something other than the satisfaction of a sensible need.

### ***The Higher Faculty of Feeling***

In the *Groundwork* (G 4:401n), Kant contrasts his characterization of respect as a feeling “self-wrought by means of a rational concept” with a description of feelings that are “received by means of influence” (i.e., result from an object affecting the senses). He mirrors this contrast in the second *Critique*, when he states

This feeling (under the name of moral feeling) is therefore produced solely by reason. It does not serve for appraising actions, and certainly not for grounding the objective moral law itself, but only as an incentive to make this law its maxim. But what name could one more suitably apply to this singular feeling which cannot be

---

<sup>55</sup> A clear statement of this position can be found at C2 5:23, “If a representation, even though it may have its seat and origin in the understanding, can determine choice only by presupposing a feeling of pleasure in the subject, its being a determining ground of choice is wholly dependent upon the nature of inner sense [interior sense], namely that this can be agreeably affected by the representation.” Kant stresses this point again at C2 5:24-5, where he states “...*pure reason* must be practical of itself and alone, that is, it must be able to determine the will by the mere form of a practical rule without presupposing any feeling and hence without any representation of the agreeable or disagreeable as the matter of the faculty of desire, which is always an empirical condition of principles” (see also C2 5:25, C2 5:62, C2 5:92 and G 4:413). The feeling being “presupposed” in these passages is clearly gratification, as indicated by his references to the inner/interior sense and to the agreeable.

compared to any pathological feeling? It is of such a peculiar kind that it seems to be at the disposal only of reason, and indeed of practical pure reason. (C2 5:76)

From these two passages, we see that Kant characterizes respect as an *a priori* feeling that is necessarily connected to pure practical reason. In other words, he thinks of respect as a higher feeling. But unfortunately, these passages do not tell us much about the nature of respect in virtue of this fact. For this, an account of the structure of the higher feelings is needed.

As a *feeling*, they will share the same basic functional structure as the lower feelings—an object is judged to fit with the subject in a way that furthers the life or activity of the subject. And in being judged to further this life, the will is determined. But from this point on, their structures diverge in several significant ways.

The primary point of divergence is in how the subject is to be understood (LM 28:248). The lower pleasures were concerned with the subject's *animal life*—the particular activity of the subject directed towards maintaining or promoting his physical/sensible well-being. Furthering that life was understood as satisfying the sensible needs stemming from his finite nature. The higher feelings, in contrast, are concerned with the subject's *spiritual* or *rational life*. This life consists of a very different sort of activity, one that can be variously described as rational, spontaneous, or free. In other words, it is a type of activity whose source of causality is other than (and so independent of) natural causality (C2 5:29, C2 5:55 and C2 5:67). Because of this, Kant also characterizes the subject of these

judgments as *freedom*.<sup>56</sup> When he does, I believe we move closer to the specific activity he had in mind—universal law-giving (C2 5:33). This law-giving activity, in turn, determines the remaining features of the higher feelings.

Since the activity of the spiritual life does not originate from the being's finite nature, it involves no sensible needs, and so no gratification. To put this another way, we do not have a sensible need to will, or even to will well, so we cannot feel the gratification of a need in the successful activity of our own willing (though we can feel self-contentment, a sensation of pleasure distinct from gratification because it is not based on a need and so is not presupposed as a condition of moral motivation) (LM 28:257-8, C2 5:38-9 and NF #7202 19:279/467). Furthering the spiritual life, then, is not to be understood as a promotion of the *well-being* of the subject, rational or otherwise.

Instead, for the spiritual life to be furthered, an object must be judged to fit with, or promote, its universal law-giving activity. Implicit in the representation of this fit must be the *form* of this activity. That is, in order for anything to further the activity of the spiritual life, it must itself qualify for, in the sense of having the proper form, a giving of universal law. However, this form is nothing other than the moral law itself (the *formal principle of universal law-giving*). So any object that is judged to fit with the form represented by the moral law will at the same time be judged to further the universal law-giving activity of the spiritual life, and so to further its freedom. In essence, the moral law is the form of the

---

<sup>56</sup> “Now if I feel that something agrees with the highest degree of freedom, thus with the spiritual life, then that pleases me. This pleasure is intellectual pleasure. One has a satisfaction with it, without its gratifying one. Such intellectual pleasure is *only in morality*...All morality is the harmony of freedom with itself. E.g., whoever lies does not agree with his freedom, because he is bound by the lie. *Whatever harmonizes with freedom agrees with the whole of life. Whatever agrees with the whole of life, pleases*” (LM 28:249-50, see also C2 5:73 and C2 5:132).

activity of the spiritual life itself (C2 5:73). Therefore, in the higher feelings, the principle (the moral law) and the subject (the spiritual life) are *necessarily* linked.

Kant gives a detailed argument for this mutual implication in chapter one of the second *Critique*. This argument is familiar enough, so I will not further explain the connection here. However, this link is important because it marks a significant change in what ultimately grounds the fit between the subject and the object. That is, with the lower feelings, underlying the fit was the relationship between the object and the subject's needs. Thus, the gratification expected to result from this fit dictated the principle. But with the higher feelings, this order is reversed. The necessary link between the subject and the principle grounds the fit. Thus, the principle dictates the objects, and does so without involving our sensibility. This reversal in concepts, where the moral principle must come before the concept of the good (the object of the higher feeling), is what Kant states is required to keep the moral principle from being based on a sensation of pleasure (gratification), and so from being empirical (C2 5:9 and C2 5:58). It is in following *this* point (the order between the principle and the concept of the good) that Kant sometimes refers to the lower pleasures as those which "precede" the moral law, and the higher pleasures as those which can only "follow" from it. With the lower feelings, the expectation of gratification *precedes* the principle in the order of determination. With the higher feelings, the feeling itself (as a judgment) follows from the principle as its result (NF #7320 19:316/478, C2 5:62 and C2 5:117).<sup>57</sup>

---

<sup>57</sup> This is in contrast to how the point is often understood, where "preceding/succeeding *the moral law*" is thought to imply preceding/succeeding *the determination of the will*, which would then make respect a

At this point, we have the subject of a higher feeling—the spiritual life, and the principle necessarily associated with it—the moral law. Together these determine which objects will fit with the universal law-giving activity of the spiritual life. But what sort of objects might these be? With the lower feelings, we saw that the objects could take a wide range of forms, from physical objects to concepts of reason. The opposite, as one might expect, holds true for the higher feelings. Clearly, physical objects or states of affairs, i.e., things that are empirically conditioned, are immediately excluded from consideration because these objects require a physical need being satisfied by their existence. Because of this need, the objects would then have to relate to the animal, rather than the spiritual, life of the subject, making it a lower feeling. So any object of pleasure that is *external* to the will itself will necessarily be an object of the lower faculty.

It follows then that the objects of the higher feelings must be internal to the activity of the will—i.e., they must be general concepts of action. But many of these concepts will not fit either because they do not contain the right form. False promising, e.g., is a concept that, when judged in relation to the freedom of more than one will, gives rise to a contradiction. In other words, it is a concept that hinders the freedom/activity of the will to which one is falsely promising. Therefore, it does not qualify as an instance of *universal* law-giving.

From the failure of false-promising, we see that the object of a higher feeling must be a concept of action that can potentially fit with, or promote, the spiritual activity of *all*

---

sensation of pleasure consequent to this determination (and so not a motive). This is the position Guyer appears to take.

rational beings. In other words, the concept of action must be judged to fit with the essence of freedom itself, and in doing so, please any being that has a source of freedom within him—whether God, angels, Saturnians, or the like.<sup>58</sup> The only concepts that can fit this condition are those of pure practical reason (the moral concepts), because they alone can qualify as a universal law (with respect to either the law’s form, or its necessary objects).<sup>59</sup> In the *Lectures on Metaphysics*, Kant gives two examples of these (moral) concepts—truthfulness and alms-giving (LM 28:253). When truthfulness, e.g., is judged to fit with the universal law-giving activity of freedom, one is, in essence, universalizing the concept of action to see if any contradictions arise. If none do, then truthfulness is judged to contain the proper form, and so would qualify as an instance of universal law. That is, the form of the moral law is recognized to be implicit in the concept itself (G 4:402 and C2 5:109-10). Because of this, when truthfulness determines the will, it increases the will’s universal law-giving activity (by being an instance of that activity), and so furthers the subject’s spiritual life (LM 29:896). As a result, truthfulness is judged to be *morally good*, the specific predicate Kant reserves for objects of the higher feelings (LM 28:248-9). In contrast, the

---

<sup>58</sup> “...the [morally] good must also please those beings who have no such sensibility like ours, but that does not hold with the agreeable and the beautiful” (LM 28:252, see also C3 5:209-10 and LM 29:892). This is why Kant states that God and holy wills are motivated by the feeling of *love* for the moral law, because even in an impassible spiritual being such as God, the faculty of desire is determined by a feeling.

<sup>59</sup> In other words, such a concept either qualifies as a giving of universal law itself, or it helps to promote this law-giving activity in other ways. An instance of the latter is the concept of humanity as an end in itself. This concept is not a representation of an action, or an object to be effected, but is instead a representation of something we should not act *against*. If we do, then our law-giving activity will fail to be universal. So in adopting humanity as an object of the will, we can be said to further its universal law-giving activity by preventing possible violations/hindrances to it (G 4:437 and C2 5:73). In addition, I believe the concept of humanity can also directly promote the activity of the will when considered as the basis for the duties to humanity. If so, then the *love of one’s neighbor* (i.e., love of humanity)—one of the four feelings in the *Metaphysics of Morals* thought to “lie at the basis of morality”—is also a higher pleasure and so a potential moral motive (MM 6:399-402 and G 4:428-9).



concept of false promising is judged to hinder the spiritual life (by suggesting an action that cannot be universalized), and so is at the same time judged to be *evil*. Because the objects of these feelings are concepts that are internal to the activity of the will itself, and so independent of any object or state of affairs that might be brought about in the world, I call these *formal*, as opposed to *material*, pleasures. That is, they are pleasures in the *form* of good willing itself (and as such, they are emotions strikingly similar to Stoic *eupatheiai*).

Now, when the higher pleasures judge a concept to be morally good, they are judging it to be *good in itself*, rather than *good for* satisfying a need. As noted earlier with the lower pleasures, the gratification of a need is the condition of an object fitting with the subject. So when I head outside for some fresh air, it is *for the sake of gratifying a need*. Hence, the fresh air is only mediately good. But a higher feeling, in being independent of any needs, must have a different condition of fit—the moral law. So when I tell the truth, it is not for the sake of some need or further end (such as garnering trust), and so not for the sake of gratification. Instead, it is *for the sake of the moral law* itself, where the moral law (i.e., qualifying for universal law-giving) is the condition. So in a word, truthfulness pleases the will in itself, and not in terms of what it promises. Therefore, it is *good in itself*, and so *immediately* good (C2 5:59, C2 5:62, C3 5:207, LM 29:891-2 and NF #1020 15:456/407-8).<sup>60</sup>

---

<sup>60</sup> In other words, whatever establishes the condition of the fit is what I believe ultimately determines the will. This is gratification with the lower pleasures, because an object that is not represented as promising gratification cannot motivate. With the higher feelings, it is the form of universal law-giving (the moral law) contained in the moral concept of truthfulness. If truthfulness did not have this form, then it too would not motivate (because it would not further the spiritual life of the subject). In both cases, the condition is contained in the concept or representation of whatever determines the will, and so is what ultimately does the motivational heavy-lifting. This relationship between the object of the pleasure and its condition for

Kant also argues that truthfulness is judged to be good in itself, or objectively good, because this judgment has a certain type of validity. Because truthfulness is judged purely in relation to the subject's rational nature, and so independently of his sensibility, this judgment represents a relationship to an object necessarily shared by all rational beings—i.e., it is a “universal judgment that has universal validity and is valid for everyone independent of the particular conditions of the subject” (LM 28:248, see also C3 5:209-10, C3 5:212-13, LM 28:252-3, LM 28:257-8, NF #711 15:315-6/495, NF #824 15:368/507 and NF #6598 19:103/420), where being independent of subjective conditions (sensibility) is a criterion for practical objective validity.<sup>61</sup> Because they are based on universally valid grounds, Kant calls these feelings *objective*. This is in contrast to the lower feelings, which he describes as *merely* subjective because they are based on the privately valid grounds of the senses, and thus hold only for the subject making the judgment.<sup>62</sup> Because the higher

---

motivating can be understood as the “what” and the “why” of the will (to follow a distinction made by St. Anselm of Canterbury in chapter 12 of *De veritate*). The object is *what* the will hopes to achieve, and the *why* is the reason for willing it (gratification or morality).

<sup>61</sup> Kant briefly mentions this criterion for objective validity in the practical domain at C2 5:21, where he states that “it is requisite to reason’s lawgiving that it should need to presuppose only *itself*, because a rule is objectively and universally valid only when it holds *without the contingent, subjective conditions that distinguish one rational being from another* (my italics).”

<sup>62</sup> Kant calls the lower feelings *merely* subjective because they are subjective in two different senses. In the primary sense, both higher and lower feelings are subjective because they are *practical*. That is, they represent the relation of an object *to a subject*, and it is in virtue of this represented relation to the subject that they motivate. But the lower feelings are also subjective because they are based on the particular sensible constitution of the subject, i.e., his contingent needs and circumstances. These feelings must have this subjective basis because their primary function is to promote the subject’s well-being (by conferring value on an object in terms of its relation to the *needs* of the subject). The higher feelings, in contrast, are *objective* in this sense because they judge a concept to agree with the activity of rational nature itself, and so to be universally pleasing to all rational beings. So unlike agreeableness, moral goodness *is* a property of the concept itself. As an objective property, when the concept is cognized by feeling (i.e., cognized in relation to the subject’s will), a determining ground is necessarily produced. Thus, we are always disposed to moral action regardless of our circumstances (and will so act as long as no hindrance from a lower feeling is

feelings are objectively valid judgments that determine the will, they have the full status of a practical cognition.

Finally, because these feelings are judgments that hold universally and necessarily for all rational beings, a moral concept such as truthfulness can be the “*very same determining ground* of the will in all cases and for all rational beings” (C2 5:25). In other words, these judgments generate the basis for morality, because they provide a universal standpoint (freedom in general) from which to view our actions. When we all judge our actions from this same standpoint, we share the same determining ground. As a consequence, our higher volitions will necessarily be in harmony with each other. And this harmony is the essence of Kantian morality (C2 5:28, NF #6621 19:114-5/425-6 and NF #7202 19:279/467).

At this point, we can now put Kant’s definition of the higher faculty of feeling into context. He states that

Objective satisfaction or dissatisfaction [higher feelings of pleasure or displeasure], or judging objects according to universally valid grounds of the power of cognition, is the higher faculty of pleasure and displeasure. This is the faculty for judging of an object whether it pleases or displeases from cognition of the understanding according to universally valid principles. If something is an object of intellectual satisfaction, then it is good; if it is an object of intellectual dissatisfaction, then it is evil. –Good is what must please everyone necessarily. (LM 28:249)

A higher feeling is thus an *a priori* judgment representing a fit between truthfulness and freedom (two *a priori* concepts of reason) according to the principle of morality (an *a priori*

---

encountered). This is part of what makes these feelings *moral*. It is in being able to characterize the higher feelings in this way, as both subjective and objective, that Kant is able to explain how the *objective* moral law can become *subjective*, i.e., how there can be a purely rational moral motive (LM 28:257-8, C2 5:73 and C3 20:245).

law). From this representation of fit, truthfulness is judged to be good in itself, and so morally good. It is thus considered to contain the form of the moral law, and so when it determines the will, so, too, does the moral law itself.

We can now come full circle and return to Kant's discussion of the higher faculties of the mind. Kant identified the higher faculties as forms of intellectuality, spontaneity or self-activity, which are the marks of our rational capacities. The higher faculty of cognition is thus reason itself. The pure will, as the higher faculty of desire, is identified with reason in its practical function. Now we can make the final identification. Since the higher feelings are universally and objectively valid, *a priori* judgments, they must be a product of reason. So the higher faculty of feeling, as their source, must be identical to a form of reason as well. But rather than specify a third function of reason, I believe Kant includes it, along with desire, under the general heading of *practical reason*, because it borrows its principle of judgment from the higher faculty of desire. To put this another way, both are forms of practical reason because both are ultimately needed in order to will at all (feeling being the means for bringing certain representations under the purview of desire). This identification of the higher faculty with a form practical reason thus makes sense of the seemingly inconsistent ways that Kant refers to the determining ground of the will—as reason itself, as the moral law, and as respect. Because respect is a higher feeling, and so is identical to an activity of reason based on the moral law, these three references are roughly equivalent.

### III. Respect

With this new account of feeling, the problem we began with has been resolved. It was previously assumed that all feelings were empirical, and so as a feeling, respect was incompatible with being a moral motive. But when respect is properly understood as a purely rational higher feeling, no such problem arises. As an action-initiating evaluative judgment, respect qualifies as both a feeling and a moral motive without violating the metaphysical constraints of Kant's theory. However, this cannot yet be the full story. Respect is the moral feeling found in *sensibly affected creatures*—creatures with both rational *and* animal natures. This is in contrast to *love* for the moral law found in holy wills and God, a feeling distinct from respect in that these beings do not have potentially wayward inclinations to control (C2 5:83-4, C2 5:32, G 4:414 and G 4:439). It thus follows that the intersection of two distinct natures in some creatures poses an additional challenge for Kant, a challenge that I believe his specific account of respect in the second *Critique* was developed to answer.

So far, the higher and lower faculties of feeling have been described as two discrete sources of motivation. Given that animals have no spiritual nature, and God has no animal nature, this would make sense. The actions of animals must be determined by the lower feelings alone, and those of God solely by the higher. Human beings, however, as finite rational creatures, possess *both* of these natures. As a consequence, two potentially independent sources of motivation are contained within a single being, and this creates a problem. If these sources function independently (i.e., as fully determining or sufficient motives), they could conflict—e.g., our lower nature might motivate us to lie to achieve

some end, while our higher nature motivates us to avoid lying, regardless of the end to be achieved. In such a case, either our agency would fragment, or our faculty of desire would contain a contradiction, both of which are deeply problematic. So in order to maintain the unity of our psychology, one or the other must emerge as the dominant, or sufficient, motive (R 6:36 and A 7:277).

Intellectualist accounts of morality, most notably the Stoics, often suggest that we should operate exclusively from our rational nature by simply suppressing our lower nature. Though Kant is often portrayed as holding this view, it is actually not an option for him. As finite creatures, we must be able to maintain our physical existence by satisfying our basic needs. So if the lower feelings were to be suppressed, then the higher feelings would have to step up and subsume this role within their general function. As it turns out, however, they are not suitable for this task. In order to satisfy a physical need, we must be able to engage in a particular action in the world—i.e., I must actually go outside for a walk in order to clear my head. But this sort of action requires the adoption of an *empirical* object or state of affairs as the matter of my will. The problem is, because the higher feelings are *formal*, they cannot provide this type of matter. Instead, they can only motivate me to adopt general principles of good-willing, such as honesty, and this activity is confined to the will itself. It thus follows that the higher feelings alone cannot promote my physical well-being. Somehow the lower feelings, as *material* feelings, must be involved. It is because of this that Kant describes them as “an unavoidable determining ground of [our] faculty of desire” (C2 5:25, see also R 6:36, R 6:58, G 4:415 and LM 29:1016).

This is where the particular problem for respect surfaces. How can a feeling of one type emerge as the sufficient motive, without also necessarily suppressing the feelings of the other type? The default answer points to the relative strength of the two feelings. The strongest feeling dominates by inhibiting only those weaker feelings in conflict with it. But Kant cannot adopt this solution for two related reasons. First, it is a purely passive, weight model of motivation whereby the strongest feeling essentially moves us to act. Morality, however, requires active motivation—i.e., we must freely determine ourselves to act (C2 5:23-25). More importantly, however, it would fail to explain how the two feelings can be properly or objectively ordered. When a concept such as honesty, as the object of a higher feeling, determines the *will* (the higher faculty), it has the superior status of an *unconditional law* governing the faculty of desire. Because of this, respect cannot simply mingle with the lower feelings, as if it were merely one among many to be compared and weighed. Instead, when it motivates moral action, it must do so in part by *governing*—i.e., by constraining and ordering our lower nature according to the moral law, thus bringing the manifold of our desires into harmony. More specifically, it must impose its objects, the moral concepts (the laws of the rational order), onto the lower feelings (sensibility) in a way that allows the morally permissible inclinations (the lower feelings and their associated desires), while excluding the wayward/antithetical ones (C2 5:65, C2 5:43, C2 5:78, C2 5:159 and G 4:395). As a result, respect cannot simply determine the will. It must also cross over into the natural order and influence choice, the *lower* faculty of desire.<sup>63</sup>

---

<sup>63</sup> “In general, nature seems to us to have in the end subordinated sensible needs for the sake of all our actions. Only it was necessary that our understanding at the same time projected universal rules, in accordance with

The question still remaining, however, is *how*. It cannot directly bring the lower feelings under its sphere of control, because the lower feelings are not concepts of reason, and thus cannot be objects of respect. So there must be some intermediary linking the two feelings. The most likely candidate is a maxim, or subjective principle of action (generated by choice in relation to an object of the lower feelings), because it is on this subjective principle that one actually comes to act. Put more simply, if respect cannot govern the lower feelings directly, the next best option is for it to guide or constrain on the basis of the principles of action they generate. So if the maxim, rather than the feeling giving rise to it, can somehow be the object of respect, then Kant has solved his problem.

To see how this might work, consider the example of my father, Dennis, a heating and air conditioning business owner. Because he has a family, he has a need to provide for them (a form of love). In response, his lower feelings judge that *increasing his income* will help to satisfy this need. When it does, he sets it as an end to be brought about. Moreover, the best way to do this, in his opinion, is to be honest with his customers. When these are taken together, they generate the maxim on which he acts, “to increase income by honest business practices.” Now, according to Kant, this maxim contains two elements—form and matter. The matter is the end to be brought about through his action—a state of affairs in which his business has increased. The form, in contrast, is an abstraction from all the particulars of my father and his circumstances contained in the maxim. The result of this

---

which we had to order, restrict, and make coherent the efforts at our happiness, so that our blind impulses will not push us now here, now there, just by chance. Since the latter commonly conflict with one another, a judgment was necessary, which with regard to all of these impulses projects rules impartially, and thus in abstraction from all inclination, through the pure will alone, which rules, valid for all actions and for all human beings, would produce the greatest harmony of a human being with himself and with others” (NF #6621 19:114-5/425-6).



abstraction is *honesty*, a general concept of action. The distinction made here between form and matter is important because each element reflects a very different way my father can value his action. The matter is valued, and so incorporated into his maxim, because it is judged to promote his well-being. This makes the lower, material feeling the empirical determining ground of choice. The form, in contrast, is valued because it is a concept of action that will promote his universal law-giving nature—i.e., it is an object of respect. So when the form is judged to be morally good, respect is incorporated into the maxim as the *a priori* determining ground of choice. Put another way, this judgment of respect is, in essence, an act of reason holding up the subjective maxim to the dictates of the objective moral law for approval (C2 5:32-34). Thus, it is the very point at which pure reason becomes practical. Because of this, Kant describes respect as “the consciousness of a *free* submission of the will to the law” (C2 5:80).<sup>64</sup>

We now see that *both* feelings are incorporated into the maxim, and so both become determining grounds of choice, because both provide reasons for willing that action. As a result, my father can act from both self-love *and* respect. However, these two motives do not have the same status within the maxim. When both are incorporated into a single maxim, an order of subordination is forced upon them. That is, because respect is recognized as the *unconditional* determining ground, it emerges as the dominant or sufficient motive, and the lower feeling is subordinated to it. So when my father acts from

---

<sup>64</sup> If, in contrast, my father had acted unreflectively on the suggestion of his lower feeling (without first getting respect’s approval for his maxim), then his action would have been merely *in accordance with* and not *for the sake of* the moral law. In other words, he would have acted honestly only because he saw it as having prudential value, without also recognizing that it had unconditional moral value.

respect, he is, in a sense, acting from a recognition that his own pursuit of happiness is conditional upon his maxim qualifying as a universal law. This makes it possible for him to pursue his morally permissible material ends while still recognizing the unconditional demand of conformity with the objective moral law. And when he does, his higher and lower natures are unified into a single principle of action, thereby preserving the integrity of his agency.<sup>65</sup>

But what happens, then, when these two feelings conflict and so cannot be incorporated into the same maxim? This is where the constraining power of respect's governance, colorfully described in the second *Critique*, becomes most evident. Consider the case of my father's competitor, Jack. He also has a family, and so has a need to provide for them. But because he likes to avoid hard work, he prefers to increase his income by overbilling his customers when he will not get caught. Taken together, these generate the maxim on which he would act, "to increase income by dishonestly overbilling when possible." As with the earlier maxim, the matter here is judged to be agreeable and so to promote his well-being. But the form of the maxim—*dishonesty*—turns out instead to be a concept that will hinder his universal law-giving nature. As a consequence, it is judged to be evil and so to be unconditionally avoided. This puts Jack in an interesting position. As with my father, he has both an empirical and an *a priori* determining ground of choice available. But since these two feelings now oppose each other, he can act on only one. If he recognizes the unconditional status of respect, then it will be the dominant motive and he

---

<sup>65</sup> "The *praxis* of morality thus consists in that formation of the inclinations and of taste which makes us capable of uniting the actions that lead to our gratification with moral principles. This is the virtuous person, consequently the one who knows how to conform his inclinations to moral principles" (NF #6619 19:113/425).

will reject the dishonest maxim. However, when the maxim is rejected, the matter—the object of the lower feeling—is rejected with it. This then leaves his original need unsatisfied, which results in the sensation of pain rather than gratification. Kant describes this pain as the sensible feeling of humiliation that follows from respect’s thwarting of inclination (C2 5:75). It is how we come to *feel* (affectively) the constraint respect puts on the power of choice. Kant makes one last point here. Whether respect positively promotes our higher nature (as with my father) or negatively constrains the wayward tendencies of our lower nature (as with Jack), it still meets Kant’s definition of a feeling. “For, whatever diminishes the hindrances to an activity is a furthering of this activity itself” (C2 5:79).

Then again, Jack could also just ignore the demands of respect and act on his dishonest maxim anyway. However, the only way for this to be possible is if Jack’s judgment were somehow impaired, making respect *appear* to have equal status with the lower feelings when it in fact does not. There are two types of error in judgment that can cause this to happen. The first is to some extent inevitable because Jack, as a human being, is a finite creature with a limited rational capacity. Because of this, his representation of a moral concept will always be, to some extent, inaccurate (C2 5:151-161 and MM 6:399-400). The less accurate it is, the less forceful respect’s authority will appear to be. This is why Kant frequently mentions that we should strive for purity in our representation of the moral law, because this is how we cultivate the moral feeling of respect (MM 6:400, C2 5:156-7, R 6:46, R 6:83, G 4:405 and G 4:410-11).

But the more serious challenge to the authority of respect can be found with the error prone lower feelings. Kant frequently notes their tendency to charm, to distort our

deliberative field, to alter our attention, and other nefarious activities. They have this power because their motivational force is due, in part, to their influence on attention and deliberation. As judgments about the agreeable, they confer value onto objects on the basis of our needs. Our attention is then directed to those objects so that the underlying need can be satisfied. The stronger the need, or the stronger the anticipation of pleasure in its satisfaction, the more valuable we judge the object to be, and so the more the object will dominate our attention or deliberative field. So when Jack faces a crisis, such as losing his home, his need to provide for his family intensifies, which in turn gives a sense of elevated importance or urgency to the end of increasing his income. When its importance begins to rival that of respect, Jack will then be tempted to make an exception to the rule “just this one time”. He still recognizes the authority of respect, but his intense fear for his family temporarily clouds or overrides this authority because all he can focus on is their well-being. When this happens, Kant calls it a mere failure of virtue (MM 6:407-8 and G 4:424). The act is immoral, but Jack himself is not yet evil.

However, if Jack’s lower feelings are left unchecked by reason, they will eventually come to assert their own unconditional status, albeit illegitimately—i.e., they will change from a form of self-love to self-conceit (C2 5:74). Because the function of a lower feeling is both to confer value and to redirect attention in relation to that value, the lower feelings must compete with each other for Jack’s attention. This competition, in turn, encourages an inflation of value, because the more attention a feeling can draw, the more likely its need will be satisfied. So in order to keep the lower feelings in check, they must be coupled with *reflection*—i.e., comparison with each other and with the demands of respect. This

reflection is important, because without it, Jack will be incapable of either happiness or worthiness to be happy (morality). More specifically, he will not be able to bring the satisfaction of his needs into harmony with each other or with the moral law. However, if the strength of a particular lower feeling succeeds in becoming excessive, it will cause Jack's attention to fixate on the particular object. And the more he fixates on the object, the more importance it appears to have. As this interplay builds, it prevents reason, even in the form of respect, from any type of reflection. This vicious cycle continues until the lower feeling sets the satisfaction of the need (or happiness in general) as the unconditional end of his action, displacing conformity with the moral law. In doing so, he elevates the status of his lower feelings above that of respect, thus reversing their proper order of subordination. Jack now does what morality requires only when it promotes his own self-interest. When this happens, Kant considers him to be in the throes of a passion, the specific type of lower feeling responsible for vice (A 7:252, A 7:265-7 and LL 24:161-7/127-32). In other words, Jack himself is now evil. So when reason is overtaken by passion, it is not overwhelmed by a blind impulse, but rather by a competing form of value judgment provided by the lower feelings—i.e., evil has a “rational origin” (R 6:41).<sup>66</sup>

---

<sup>66</sup> “The law rather imposes itself on him irresistibly, because of his moral predisposition; and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e., he would be morally good. He is, however, also dependent on the incentives of his sensuous nature because of his equally innocent natural predisposition, and he incorporates them too into his maxim (according to the subjective principle of self-love). . . . Hence the difference, whether the human being is good or evil, must not lie in the difference between the incentives that he incorporates into his maxims (not in the material of the maxim) but in their *subordination* (in the form of the maxim): *which of the two he makes the condition of the other*. It follows that the human being (even the best) is evil only because he reverses the moral order of his incentives in incorporating them into his maxims. He indeed incorporates the moral law into those maxims, together with the law of self-love; since, however, he realizes that the two cannot stand on an equal footing, but one must be subordinated to the other as its supreme condition, he makes the incentive of self-love and their inclinations the condition of compliance with the moral

Once we understand the natural tendency of the lower feelings to overestimate their own status, it becomes clear why Kant promotes a sort of *moral apathy*—a moderation of the lower feelings that prevent them from rising to the level of a passion (C3 20:196, MM 6:408-9 and A 7:253-4). In other words, he does not promote the suppression of all emotion stemming from our lower nature, as the Stoics do. Instead, he follows Augustine in focusing on only those emotions of our lower nature that “shut out the sovereignty of reason” (A 7:251, see also *De civitate Dei* XIV.9). The end of moral apathy, then, is to keep the influence of the lower feelings in check so that respect always remains in control. And keeping respect in control is, I believe, the basis of Kantian virtue (MM 6: 394, MM 6:408-9 and C2 5:84).

In the end, we clearly see that respect is indeed one of the linchpins of Kant’s moral theory. It is the point at which pure reason becomes practical, and so the point at which morality becomes possible at all. Kant emphatically describes this importance when he states that respect is “an estimation of a worth that far outweighs any worth of what is recommended by inclination, and that the necessity of my action from *pure* respect for the practical law is what constitutes duty, to which every other motive must give way because it is the condition of a will good *in itself*, the worth of which surpasses all else” (G 4:403).

---

law” (R 6:36, see also R 6:21 and R 6:41). Kant describes the principle of such a will as “*Love yourself above all, but God and your neighbor for your own sake*” (C2 5:83n).

## BIBLIOGRAPHY

- Anselm, Saint, Archbishop of Canterbury, 2007, *Anselm: Basic Writings*, trans. Thomas Williams, Indianapolis: Hackett Publishing Company.
- Augustine, Saint, Bishop of Hippo, 1998, *The City of God against the Pagans*, ed. and trans. R. W. Dyson, Cambridge: Cambridge University Press.
- Kant, Immanuel, 2006, *Anthropology from a Pragmatic Point of View*, ed. and trans. Robert B. Loudon, Cambridge: Cambridge University Press.
- , 2000, *Critique of the Power of Judgment (C3)*, ed. and trans. Paul Guyer and trans. Eric Matthews, Cambridge: Cambridge University Press.
- , 1997, *Critique of Practical Reason (C2)*, ed. and trans. Mary Gregor, Cambridge: Cambridge University Press.
- , 1998, *Critique of Pure Reason (C1)*, ed. and trans. Paul Guyer and Allen Wood, Cambridge: Cambridge University Press.
- , 1997, *Groundwork of the Metaphysics of Morals (G)*, ed. and trans. Mary Gregor, Cambridge: Cambridge University Press.
- , 1983, “Idea for a Universal History with a Cosmopolitan Intent” (IUH) in *Perpetual Peace and Other Essays*, Indianapolis: Hackett Publishing Co.
- , 2012, *Lectures on Anthropology (LA)*, ed. and trans. Allen Wood and Robert Loudon, trans. Robert Clewis and G. Felicitas Munzel, Cambridge: Cambridge University Press.
- , 1997, *Lectures on Ethics (LE)*, ed. and trans. Peter Heath and ed. J. B. Schneewind, Cambridge: Cambridge University Press.
- , 1992, *Lectures on Logic (LL)*, ed. and trans. J. Michael Young, Cambridge: Cambridge University Press.
- , 1997, *Lectures on Metaphysics (LM)*, ed. and trans. Karl Ameriks and Steve Naragon, Cambridge: Cambridge University Press.
- , 1996, “Lectures on the Philosophical Doctrine of Religion” (LPDR) in *Religion and Rational Theology*, ed. and trans. Allen Wood and George di Giovanni, Cambridge: Cambridge University Press.

- , 1996, *Metaphysics of Morals* (MM), ed. and trans. Mary Gregor, Cambridge: Cambridge University Press.
- , 2005, *Notes and Fragments* (NF), ed. and trans. Paul Guyer and trans. Curtis Bowman and Frederick Rauscher, Cambridge: Cambridge University Press.
- , 1998, *Religion within the Boundaries of Mere Reason* (R), ed. and trans. Allen Wood and George di Giovanni, Cambridge: Cambridge University Press.
- , 1983, “Speculative Beginning of Human History” (SBHH) in *Perpetual Peace and Other Essays*, Indianapolis: Hackett Publishing Co.
- Knuuttila, Simo, 2004, *Emotions in Ancient and Medieval Philosophy*, Oxford: Oxford University Press.
- Graver, Margaret, 2007, *Stoicism and Emotion*, Chicago: University of Chicago Press.
- Guyer, Paul, 2000, *Kant on Freedom, Law, and Happiness*, New York: Cambridge University Press.
- Reath, Andrews, 1989, “Kant’s Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination,” *Kant-Studien*, 80: 284-302.
- Schapiro, Tamar, 2011, “Foregrounding Desire: A Defense of Kant’s Incorporation Thesis,” *Journal of Ethics*, 15: 147-167.