

KANT'S THEORY OF EMOTION:
TOWARD A SYSTEMATIC RECONSTRUCTION

Uri Eran

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the Department of Philosophy,

Indiana University

January 2021

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.

Doctoral Committee

Allen W. Wood, Ph.D. (Chair)

Marcia Baron, Ph.D.

Sandra L. Shapshay, Ph.D.

Johannes Türk, Ph.D.

January 4th, 2021

Copyright © 2021

Uri Eran

In memory of my beloved father, Oded.

Acknowledgments

Philosophical work, like any other, is never the product of a single person. The present dissertation is no exception. In thinking and writing about Kant's theory of emotion I consider myself very fortunate to have had the opportunity to converse - be it in person, in email correspondence, or in my own mind while reading and thinking of their work - with an excellent group of philosophers. Allen Wood, who supervised this project, is famous in the philosophical community for his outstanding work on Kant, as well as for his unrelenting commitment and availability to students. What is perhaps less known is that Allen can also be flexible and open-minded in considering interpretations other than his own. This open-mindedness, along with Allen's courage in giving voice to unorthodox views in the history of philosophy, and his insistence that good interpretive work requires "keeping your eye on the ball" of the philosophical issues at stake have been a source of inspiration for me, and made this project a better one. For setting forth these models of virtues, and for the countless hours he has put into reading and commenting on my work, I cannot thank him enough.

Working with a scholar of Allen's rank, who has written on virtually every aspect of Kant's philosophy, has the drawback that one is constantly tempted to repeat the teacher's doctrines rather than "think for oneself" and come up with a new perspective. Luckily for me, Marcia Baron was there to make sure I resisted this temptation (which Allen, a devout advocate of the Enlightenment, would too surely have wanted me to resist). Whenever I would take some interpretation for granted, ignore a passage that threatened my reading, gloss over some distinction, or simply avoid a difficulty that needs to be addressed, it was Marcia's voice - either the actual one, or its written or imagined incarnation - that had set me straight. Marcia's high standards of rigor, clarity, and precision probably remain regulative ideas for me rather than

goals I have attained. But for always pushing me to strive for them, as well as for her genuine care and emotional support in times when philosophical inquiry seemed to make little sense, I am deeply grateful to her.

Sandra Shapshay's exciting courses on Schopenhauer and Nietzsche, as well as on Kant's third *Critique* have been another source of significant influence on this project. Sandy's generous, exuberant and kind character was often a breath of fresh air in the barren landscape of academic life. Her moral support and optimistic spirit were also indispensable and much needed during my time at IU.

In addition to the members of my dissertation committee from the philosophy department, I am also grateful to Johannes Türk, who has served on my committee as the representative of the PhD minor in History of German Thought. Johannes' course on the modern novel had taught me the value of a more tolerant, non-moralizing approach to the emotions. Fellow philosophy graduate students at IU - in particular Sean Murphy, Dan Buckley, Andrew Smith, Ivan Verano and Levi Tenen - were always available and excited to talk philosophy, and provided an excellent intellectual community, with a rare combination of outstanding philosophical skills on the one hand, and a modest and supportive spirit on the other hand. Helpful comments and feedback were also provided by Dina Emundts and Tobias Rosefeldt and participants in their colloquia at FU and HU Berlin (respectively), by Dieter Schönecker and participants in his colloquium at Universität Siegen, as well as by Claudia Blöser, Janelle DeWitt, Ido Geiger, Leonard Katz and Angela Smith.

A foreign exchange program at FU Berlin, made possible by Indiana University's Office of the Vice President for International Affairs, allowed me to improve my German significantly, invest more time in research, familiarize myself with the European Kant scene, and get valuable

feedback from participants in philosophy colloquia in Berlin and Siegen. And a Minerva Stiftung Fellowship, hosted by Dina Emundts at FU, has given me ample time to improve this work and publish parts of it in journals.

For their support, friendship and feedback, I also wish to thank Eyal Akiva, Yael Atia, Lucía Cores Sarría, Dikla Bytner, Eran Dorfman, Yiftah Elazar, Dan and Gili Eran and their families, Gal Eshel, Maria Fink, Naveh Frumer, Noam Hoffer, Antonios Kalatzis, Norma Musih and Dani Szpruch, Ori Rotlevy, David Tal, Doron Tirosh, and Noam Yuran and his family.

Last but not least, I wish to thank my parents, Daphna and Oded, whose care, love and support have enabled and encouraged me to devote myself to the non-lucrative life of philosophical research. In the middle of 2020, at the age of 65, we lost Oded after a short and brutal fight with cancer. This dissertation is dedicated to him, my first teacher in philosophy. When I once asked my father why he insisted on gradually replacing the different parts of the old Parker pen he got for his Bar Mitzvah rather than get a new one, he replied: because it wouldn't be the same pen. And to top it off, he went on to compare his Parker with the human body, that miraculously preserves its identity despite replacing its cells every seven years. A lover of all things intellectual, whose analytic skills are every philosopher's dream, my father was a brilliant man by any standard. Unfortunately, as often is the case with cerebral folk, although a genuinely loving and caring person, my old man was not great in recognizing and expressing his emotions. A dismissive attitude toward this aspect of the mental life was apparent in his frequent use of the word "pathetic" to refer to behavior he regarded as inappropriate or immature. But in spite of this stern disposition, his love and affection were always felt and in his last days, he could no longer afford to repress. Neither could I. I can only hope that my engagement with emotions in Kant is a first step toward reconciling *logos* and *pathos*, in Kant, and in myself.

Uri Eran

KANT'S THEORY OF EMOTION: TOWARD A SYSTEMATIC RECONSTRUCTION

Putting together Kant's theory of emotion is complicated by two facts: (1) Kant has no term which is an obvious equivalent of "emotion" as used in contemporary English; (2) theorists disagree about what emotions are. These obstacles notwithstanding, my dissertation aims to provide the foundation for a reconstruction of Kant's theory of emotion that is both historically accurate and responsive to contemporary philosophical concerns. In contrast to available approaches which rest on contested assumptions about emotions, I start from the generally accepted and reasonable premise that what we call "emotions" refers in Kant to a set of mental states, some of which he associates with the feeling of pleasure and displeasure ("feelings"), others with the faculty of desire ("desires"). I then proceed to examine the nature of these two kinds of mental states and their proper treatment. I argue that Kantian feelings are representations of objects' relation to the subject, that have a felt quality, and dispose their subject to certain behaviors. While feelings can only motivate action by causing desires and have no temporal direction, desires - except for certain wishes - are future-directed, which allows them to motivate actions immediately (but they need not bring action about). Equipped with this account of feelings and desires, I proceed to examine the kind of treatment Kant prescribes for them, and argue that feelings (except affects) should be cultivated, that is, acquired and improved so that they could be used to pursue rational ends, while inclinations, i.e., habitual sensuous desires, should generally be disciplined, that is, constrained by rules. The resultant picture is compelling because it rests on minimal assumptions about emotions and successfully incorporates the phenomenological, evaluative, and dispositional functions traditionally associated with emotions.

Table of Contents

Dedication	iv
Acknowledgments	v
Abstract	viii
Table of Contents	ix
Introduction: From Our Emotions to Kant's	1
Chapter 1. How to Approach Emotions in Kant	9
1.1. Where is Kant's Theory of Emotion?.....	9
1.2. What's the Problem?.....	11
1.3. Frierson's Denial of a General Theory.....	15
1.4. Cohen's Account of Emotions as Feelings.....	23
1.5. Deimling's "Pragmatic Concept of Emotions".....	27
1.6. Emotions as a Set of Feelings and Desires.....	34
Chapter 2. More than A Feeling: Kant's Tripartite Account of Pleasure	38
2.1. Phenomenological Qualities, Dispositions or Evaluations?.....	38
2.2. Pleasure, Satisfaction, and Feeling.....	41
2.3. Pleasure as a Subjective Representation and a Felt Quality.....	46
2.4. Pleasure as a Disposition.....	54
2.5. Two Alternative Accounts.....	59
2.6. Lingering Objections.....	65
Chapter 3. Kantian Desires: A Holistic Account	71
3.1. Between Pleasure and Action.....	71
3.2. Desire and Its Species.....	74

3.3. Desire and Foresight.....	79
3.4. Desires as Dispositions.....	82
3.5. Desire and Pleasure.....	90
3.6. Possible Objections.....	99
3.7. A "Motivation First" Holistic Theory of Desire.....	102
Chapter 4. Which Emotions Should Kantians Cultivate (and Which Ones Should They Discipline)?.....	105
4.1. Should a Good Kantian Cultivate Her Emotions?.....	105
4.2. Discipline and Cultivation.....	107
4.3. Cultivate Feelings, Discipline Inclinations.....	113
4.4. Why Different Treatments?	129
4.5. Defending the Cultivation of Feelings.....	134
Conclusion. Two Theories of Emotion?.....	140
Note on Translations and Abbreviations.....	148
Bibliography.....	151
Curriculum Vitae	

Introduction: From Our Emotions to Kant's

"Kantian emotions" is a hot topic. In the last three decades or so, with the rise of virtue ethics and the revival of general-philosophical interest in the emotions, Kant scholars have done considerable work that aims to uncover the emotional underpinnings of Kant's moral theory and of his theory of action and agency. The first wave of this trend can be traced back to the 1990's works of Barbara Herman (1993, Chapters 1 and 4), Marcia Baron (1995, Chapter 6), Nancy Sherman (1997, Chapter 4), and Allen Wood (1999, 150-83). Against the familiar caricature of Kant as the "cold sage of Königsberg" who regards emotions as blind, non-rational forces at best and as a serious threat to morality at worst, these scholars and others have convincingly argued that emotions are not merely an ineliminable part of Kant's moral theory, but are also indispensable for leading a virtuous life.¹ A "second wave" soon followed, starting in the early 2000's, when scholars began to shift their attention from the place of emotions in Kant's ethics to the more fundamental issue of the very nature of emotions in Kant (e.g., Sorensen 2002; Borges 2004; Deimling 2014, Frierson 2014a; Williamson 2015; Cohen 2017, 2020; DeWitt 2018).

In a recent helpful summary of the fruits of these efforts, Alix Cohen distinguishes between five different models of emotions in Kant:

1. The pain model: Cohen attributes this model to John Sabini and Maury Silver who had argued a few years prior to the aforementioned "first wave", that "Emotions [for Kant] are brute forces unconnected with higher mental functions. Pain is the obvious model. Pain is a brute force; it is beyond the will; it is, or at least typically is, independent of reason." (Sabini and Silver 1987, 166, as quoted in Cohen 2017, 665). Similarly, Cohen says, Paul Guyer defines emotions in Kant

¹ This trend has been challenged, however, by commentators who remain skeptical that Kant's attitude toward emotions is as positive as work by scholars of the "first wave" suggests, e.g., Borges (2008, 2019), Thomason (2017).

as "brute sensation-like states that are opaque in the sense that they provide no insight into their causal history." (Cohen 2017, 665).²

2. The rational model: Cohen attributes this model to Marcia Baron, who argues that "Kant does *not* hold that we are passive with respect to our emotions and feelings... It is a serious mistake to think that Kant's psychology of emotion even approximately fits the model of pain." (Baron 1995, 195, as quoted in Cohen 2017, 666). Against the pain model, Baron suggests that "it would be more plausible to criticize Kant for attributing to us too much responsibility for our feelings and emotions than to attribute to him the position that we are not responsible for them." (Baron 1995, 197, as quoted in Cohen 2017, 666).

3. The pragmatic model: Wiebke Deimling is held by Cohen as a proponent of this model. According to Cohen, Deimling suggests that emotional states can only be unified "in that they are of special interest for pragmatically and morally rational action." (Deimling 2014, 108, as quoted in Cohen 2017, 666).

4. The motivational model: Janelle DeWitt is named by Cohen as committed to this model, on which Kantian emotions function as "action-initiating evaluative judgments." (DeWitt 2014, 33, as quoted in Cohen 2017, 666).

5. The mixed model: Cohen assigns this model to Maria Borges and to Patrick Frierson. Although she does not say in what sense emotions are "mixed" according to this model, Cohen's criticism of it and her citations from Borges and Frierson indicate that they are mixed in that the realm of emotions in Kant is too heterogenous to be explained by a single model, and perhaps also in that each emotion is constituted by a mixture of several essentially different phenomena

² Guyer's account in the pages that Cohen refers to is actually of *pleasure* in Kant, not of emotions (Guyer 1997, 103-5).

(e.g., beliefs, desires, etc.).³ Accordingly, Borges argues that "Kant presents us with a very colorful, wide range of emotions, which cannot be captured in one model type"; and she adds that Kantian emotions are "intentional states as well as feelings... They have a propositional content, and also bring evaluation and cognitive elements... Beliefs and desires are constitutive of these mental states." (Borges 2004, 143, 152, as quoted in Cohen 2017, *ibid.*). Similarly, Frierson argues that "so-called 'emotions' are cognitive, affective, and volitional." (Frierson 2014a, 172, as quoted in Cohen 2017, 666).

When I started working on this project, I was baffled by the diversity of interpretive views on Kantian emotions, which Cohen's list nicely captures. Granted, some of these models are compatible with others, which suggests that the disagreement does not run very deep. For example, one could say with the rational model that we are responsible for our emotions in some ways, and also hold, as the motivational model does, that emotions are action-initiating evaluative judgments. But some models are incompatible with others, and provide very different pictures of Kantian emotions. If, for instance, emotions are brute sensation-like states (the pain model), they cannot be evaluative judgements (the motivational model); and if they cannot be captured by one model (as the mixed model says), then both the pain and the rational model seem to be inaccurate, insofar as they present the realm of Kantian emotions as more homogeneous than it actually is.

³ Cohen argues that the mixed-model "takes the apparent diversity of affective states at face value and chooses to emphasize it on the basis that the notion seems to cover numerous, distinct states that include feeling, affect, passion, inclination, desire, and even instinct." (Cohen 2014, 667; cf. Cohen 2020, 432). This suggests that on Cohen's understanding, the mixed model (mistakenly) regards the realm of emotions as too heterogeneous to be captured by a single model. It is not clear that she sees that Frierson's and Borges' work suggests that emotions are mixed in the other abovementioned sense, namely, that each emotion is a mixture of several essentially different phenomena.

It appears, then, that after three decades of concentrated effort in this field, there is still genuine disagreement on the nature of Kantian emotions. And while some of it may be due to the fact that Kant has no systematic account of emotions, there is nonetheless plenty of pertinent evidence to go on in his major ethical writings, in the third *Critique*, in *Anth.*, and in student notes from his lectures on anthropology, metaphysics, and ethics. So how come commentators provide such different accounts of Kantian emotions?

My puzzlement about the apparent disagreement started to wear off when I noticed that there are two major obstacles to providing an account of emotions in Kant. First, Kant does not have a German term that is an obvious equivalent of "emotion" as used in English nowadays. What he does have is a wide variety of terms which look like good candidates to be included in a theory of emotion, such as "feeling", "affect", "passion" and "inclination", as well as terms for more specific mental states we nowadays would probably call "emotions" (such as "love", "anger", "fear", and "envy"). There is, however, no overarching, generic term in Kant that applies to all the mental states designated by these terms; and the mental states designated by them are associated by Kant with two different mental faculties, that is, with the faculty of desire and the feeling of pleasure and displeasure. The second reason that making sense of emotions in Kant is difficult is that theorists disagree about the nature of emotions, construing them as feelings (e.g., James 1884), motivations (Scarantino 2014), evaluations (Solomon 1980; Nussbaum 2001), or some combination of these phenomena (Goldie 2000; Helm 2009; Deonna and Teroni 2012). Taken together, these two facts make it difficult to decide where in Kant we should be looking to find out how he understands emotions. This does not yet explain the disagreement on Kantian emotions, but it does suggest that a decision regarding where to look for Kant's theory cannot be based solely on anything Kant had to say about emotions. Indeed, if

what I argue in the following pages is correct, much of the disagreement about the nature of Kantian emotions and the role they ought to play in the virtuous life stems not from textual disagreement, but rather from different presuppositions about the nature of emotions and about which terms designate emotions.

These difficulties notwithstanding, my aim in what follows is to provide the foundation for a reconstruction of Kant's theory of emotion that depends on minimal, reasonable and explicit assumptions about emotions. By doing so, I hope to pave the way to a reconstruction that is faithful to the text on the one hand, and can teach us something of value about emotions on the other hand. My proposed starting point is the premise that what we call "emotions" refers to a set of mental states in Kant, some of which he associates with the faculty of desire ("desires"), others with the feeling of pleasure and displeasure ("feelings" or "pleasure and displeasure"). The mental states of the feeling of pleasure and displeasure, I will argue, combine evaluative, phenomenological and dispositional aspects; the mental states of the faculty of desire are necessarily connected with those of the faculty of feeling, and are essentially dispositions or motivations to action. Although both kinds of mental states function as dispositions, desires that are not mere wishes are directed at *future* objects and so they dispose us to action immediately, while feelings can only dispose to action by means of desires because they are directed at the subject's affective state. Equipped with this understanding of the nature of feelings and desires, I turn in the final chapter to the treatments Kant prescribes for them.

I begin in Chapter 1 by tackling the methodological problems that arise when trying to reconstruct Kant's understanding of the mental states we call "emotions". I argue that such efforts must start from assumptions about the nature of emotions or about which terms designate emotions, that determine where in Kant we will be looking when reconstructing his theory. I

proceed to examine the strategies employed by three commentators (Frierson, Deimling, and Cohen) in approaching emotions in Kant, and bring out their tacit assumptions. I then argue that we should begin our reconstruction from the generally accepted and reasonable premise that what we call "emotions" refers to a set of mental states in Kant, some of which he associates with the faculty of desire ("desires"), others with the feeling of pleasure and displeasure ("feelings"). Although it is possible that some of their species are *not* emotions on some conceptions of emotion, getting clear on the nature of feelings and desires is nonetheless a necessary first step toward reconstructing Kant's theory on the basis of minimal and generally accepted assumptions.

I proceed to examine the nature of feelings in Chapter 2. Starting with Kant's comments on the subjective, non-cognitive nature of feelings of pleasure and displeasure, I argue that they should be read such that feelings involve representations of objects' relations to the subject, and felt qualities. There is also evidence, however, that pleasures involve behavioral dispositions, and I propose that Kant's shifting between these three characterizations is best interpreted by attributing to him a tripartite account, on which feeling is an evaluation of an object's agreement or disagreement with the subject's ability or intention to act, that has a felt phenomenological quality, and disposes the subject to maintain or change its state, or to act in order to produce or not to produce an object. Such an account is both supported by the text and philosophically attractive because it incorporates the three functions traditionally associated with pleasure.

Chapter 3 examines the nature of the mental states Kant associates with the faculty of desire, i.e., desires. I first adduce evidence that desires (except for certain wishes) are directed at future objects, and that they function as motivations or dispositions to action, which can but need not bring it about. I then look at passages where Kant speaks of desire as connected with pleasure and as a pleasure in an object's existence, and I propose to read them such that desires are

causally connected with practical feelings of pleasure, but differ from them in that the former are usually future-directed while the latter are not (even when connected with representations of future objects). It is the future-directedness of desires that are not wishes that allows them to motivate actions immediately. I end by responding to some objections, and noting that this account anticipates contemporary holistic theories of desire that combine dispositional, evaluative, and hedonic components.

While the bulk of this study is dedicated to the nature of emotions in Kant, the final chapter makes use of this account in order to examine the kind of treatment Kant thinks our emotions call for. Here, I propose, Kant draws an important yet neglected distinction between the appropriate treatment for feelings and for inclinations. Feelings other than affects require cultivation, i.e., acquiring and improving them for the purpose of pursuing rational ends, while inclinations generally require discipline, i.e., constraint by means of rules. After explaining how cultivation and discipline work and differ, I propose that Kant's prescription makes sense because discipline involves preventing a mental state from motivating a forbidden action (or from motivating the omission of a required one), and inclinations can but feelings cannot motivate action immediately. Although I read Kant as prescribing that we cultivate feelings, I argue that my reading is not vulnerable to Thomason's recent objections to emotional cultivation.

The result of these four chapters is a foundation for a rich and complex account, on which Kantian emotions come in two kinds: feelings, and desires. Emotions that are feelings are evaluations of objects' relations to the subject, that have a felt quality and dispose us to action by means of desires; emotions that are desires are dispositions to action that need not bring it about, and are usually directed at future objects. It is the future-directedness of desires that are not wishes that allows them to motivate action immediately. If some Kantian feelings or desires are

not emotions, then this account provides the necessary but insufficient conditions on a mental state's being a Kantian emotion. The work done here nevertheless provides the basis from which we can infer what Kant's theory would look like by adding certain extra-textual premises about emotions (that is, premises about emotions not based on evidence from Kant's writings). But as it stands, this work has the merit that it does not rest on contested assumptions about emotions that may be alien to Kant, and can be accepted by all who agree that what we call "emotions" refers to some set of feelings and desires in Kant. And although the fact that emotions, on this account, are associated with two mental faculties suggests that they are of two natural or theoretical kinds, this need not count against its philosophical plausibility nor against its historical accuracy.

Before we begin, a few preliminary terminological remarks are in order. First, although I will sometimes be using the phrase "Kant's *theory* of emotion" (in the singular), this is not to exclude the possibility that Kant has more than one theory of the set of mental states we call "emotions", such that this set includes more than one theoretical kind.⁴ Second, unless indicated otherwise, "A Kantian X" means in what follows "X as understood in a way that best captures what Kant had written and taught about X", not "X understood in a way which is influenced by Kant". Third and lastly, I will be using here the terms "desire", "satisfaction", and "pleasure" to refer both to their "positive" and "negative" types and tokens. To indicate that I am using them to refer solely to their "positive" instances, I will either attach to these terms their "negative" counterparts ("desire and aversion", "satisfaction and dissatisfaction", "pleasure and displeasure"), or I will append to them a plus sign, e.g., "desire (+)".

⁴ The dissertation's title is therefore a bit misleading, insofar as it suggests that Kantian emotions are of a single theoretical kind. "Kant's *Theories* of Emotion" is less elegant, however.

Chapter 1. How to Approach Emotions in Kant

1.1. Where is Kant's Theory of Emotion?

As mentioned in the Introduction, the fruit of the recent work on emotions in Kant is an abundance of Kantian theories of emotion that seem to be at odds with one another. Cohen (2020) and Williamson (2015), for example, argue that Kant conceives of emotions as feelings. Deimling, in contrast, proposes that Kant has a concept of similar scope to that of our contemporary concept of emotion, that "cuts across different kinds of affective states that Kant distinguishes, such as 'feelings', 'desires', 'affects', 'inclinations' and 'passions'" (Deimling 2014, 108). Others deny that Kant has a single or general theory of emotion. Borges, for example, argues that the wide range of emotions in Kant "cannot be captured in one model type" (Borges 2004, 143); and Frierson maintains that Kant "does not have a general theory of 'emotions'." (Frierson 2014a, 168).

Why is it that commentators seem to disagree about emotions in Kant? Could the dispute be settled by appeal to textual evidence? And if not, is there still a way to make progress in our understanding of Kant, in a way that could also benefit contemporary theories of emotion? The aim of the present chapter is to answer these questions. The reason that commentators reach different conclusions regarding Kantian emotions, I argue, is that they are operating with different assumptions about what emotions are, or about which terms designate emotions. We can, however, make progress in our understanding of Kant and read him in a philosophically fruitful way if we

begin our reconstruction by looking at the nature of the two genera which all the mental states⁵ we call “emotions” are species of, namely, “feeling” and “desire”.⁶

I start by explaining the difficulties in reconstructing Kant's theory of emotion, and argue that such reconstructions must start from assumptions about the nature of emotions or about which terms designate emotions, that determine where in Kant's writings we will be looking when reconstructing his theory (1.2). I proceed to examine some strategies that commentators employ in approaching emotions in Kant, and bring out their tacit assumptions: Frierson's denial that Kant has a general theory of emotions presupposes that certain terms designate emotions in ordinary English, while others do not (1.3); Cohen's account of Kantian emotions as feelings depends on the assumption that we would call certain terms in Kant “emotions” (1.4); and Deimling's account of Kant's pragmatic concept of emotions depends on the assumption that emotions track complex values and are amenable to mediate control (Section 1.5). I then argue that we should begin our

⁵ I will be using the terms “mental state” and “mental representation” (or simply “representation”) as roughly equivalent, to refer to something that is present to the mind. I take it that this is what Kant means when he says that representations are “inner determinations of our mind in this or that temporal relation.” (A197/B242). In calling the states of the faculty of desire and feeling “mental states” I do not mean to deny that they are related to bodily states. My point is merely that as products of *mental* faculties (the faculty of desire or the feeling of pleasure and displeasure), they are mental phenomena. For readings that stresses the physiological aspects of Kantian emotions, see Borges (2008) and Williamson (2015).

⁶ As we shall see in Chapter 2, although he usually uses “feeling” to refer to a mental faculty rather than a mental state, Kant also uses it as the genus to which all mental states of the faculty of feeling belong. Regarding the term “the faculty of feeling”: although Kant sometimes lists it as one of the three primary mental faculties (*CPJ* 5:198, *FI* 20:245), at other times he reserves the term “susceptibility” (*Empfänglichkeit*, *MM* 6:212, *FI* 20:207, *Anth.* 7:153; or “receptivity” *Receptivität*, 5:58) for “the feeling of pleasure and displeasure”. According to *Anth.*, a faculty is active while a susceptibility is passive in respect to the state of the mind's representations (7:140). If the feeling of pleasure and displeasure is a susceptibility and not a faculty, it is passive with respect to the mind's representations. However, Kant also argues that “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.” (*LM* 28:228-9). This suggests that when pleasure and displeasure depend on objects' affection, the feeling of pleasure and displeasure functions as a passive *susceptibility*. In contrast, when pleasure and displeasure are independent of objects, the feeling of pleasure and displeasure functions as an active *faculty*. In what follows, I will sometimes speak of the feeling of pleasure and displeasure as a mental faculty, but this is meant to refer to this mental capacity both in its active and passive guise (i.e., both as a faculty and as a susceptibility).

reconstruction from the generally accepted and reasonable premise that what we call “emotions” refers to a set of mental states in Kant, some of which he associates with the faculty of desire, others with the feeling of pleasure and displeasure. Rather than reconstructing Kant's theory on the basis of his accounts of the different species of these mental states, which would require introducing contested assumptions, I propose we should look at what Kant says about the two genera to which they all belong, i.e., “feeling” and “desire”.⁷ Although it is possible that some of their species are not emotions, getting clear on the nature of feelings and desires is nonetheless a necessary first step toward reconstructing Kant's theory of emotion on the basis of minimal and generally accepted assumptions (1.6).

1.2. What's the Problem?

In the Introduction, I mentioned two obstacles on the way to making sense of emotions in Kant. I now want to present them in more detail, and explain just how they complicate our task. Recall the two obstacles: (1) Kant does not use any German term that is an obvious equivalent of the contemporary English term “emotion”. Instead, he uses a wide variety of terms which look like good candidates to be included in a theory of emotion, but are associated by Kant with different mental faculties. For example, Kant associates “feeling”, “affect” and “anger” with the feeling of pleasure and displeasure, while he associates “desire”, “inclination”, “passion”, and “hatred” with

⁷ Kant sometimes uses “desire” (*Begierde*) to refer to sensuous mental states of the faculty of desire, for example, when he writes: “Desiring (*Das Begehren*) *appetitio* and desire (*die Begierde*) *concupiscentia* are not one and the same. One is the *genus* and can therefore be completely intellectual, the other a *species* which is always sensuous...” (23:262, translation mine). But at other times, Kant allows that desire (*Begierde*) could be intellectual (e.g., “All desires (*Alle Begierden*) are... intellectual or sensitive.” *LM* 29:895). In what follows, when I speak of desire as the genus to which all mental states of the faculty of desire belong, I am using it in this latter, broad sense.

the faculty of desire (*Anth.* 7:251-2, 266);⁸ (2) theorists disagree not only on the nature of emotions - construing them as feelings, evaluations, motivations, or some combination of these phenomena - but also on whether they even have a common nature, such that the term “emotion” picks out a natural kind.⁹ If (1) were true but not (2), we could have reconstructed Kant's theory of emotion by assuming a non-controversial conception of emotion, and then reconstructed Kant's theory on the basis of what he says about mental states that fit this conception. If (2) were true but not (1), we could have reconstructed Kant's theory of emotion on the basis of what he says about the German term that is Kant's equivalent to our “emotion”. Taken together, however, these two facts make it difficult to decide which terms in Kant's writings we should be looking at when reconstructing his view of the mental states we call “emotions”, and whether Kant even has a single and general theory of emotions.¹⁰

I see two possible ways to reconstruct Kant's theory of emotion in light of these difficulties: (a) we could start from certain assumptions about the nature of emotions, and reconstruct Kant's theory on the basis of what he says about mental phenomena that he characterizes in a way that fits our understanding of emotions; (b) we could start from some assumptions about which terms designate emotions in ordinary English, and proceed by reconstructing Kant's theory on the basis

⁸ The absence of the term “emotion” (or a German equivalent) in Kant is not unusual for his era: as Schmitter observes, “Seventeenth century philosophers favored talk of ‘passion’ and ‘affect,’ while their eighteenth century counterparts made increasing use of ‘sentiment.’” (Schmitter 2010, §1.1). This practice is probably due to the fact, noted by Dixon, that although the word “emotion” (imported into English from the French *émotion*) was in use in the 17th and 18th centuries, “it did not become established as the name for a category of mental states that might be systematically studied until the mid-19th century.” (Dixon 2012, 338).

⁹ According to Scarantino, the three main traditions in the study of emotion are the feeling, motivational, and evaluative tradition (Scarantino 2016, 5). But the heterogeneity within the domain referred to by the generic folk-category “emotion” has led some to argue that this term does not designate a natural kind (e.g., Rorty 2004; Griffiths 1997; Kagan 2007).

¹⁰ The absence of an obvious equivalent of “emotion” in Kant is noted in different ways by Williamson (2015, 2), Sorensen (2002, 110), Borges (2004, 143), Cohen (2020, 429) and Thomason (2017, 458, fn. 2). Deimling refers to two problems similar to those mentioned above (Deimling 2014, 108,112).

of his characterizations of these terms and the phenomena designated by them.¹¹ (a) is problematic because theorists disagree about the nature of emotions, and so if we apply this method, our reconstructed theory will not be a theory of emotion at all according to some theorists. For instance, suppose we assume that emotions are evaluations of a certain kind, and reconstruct Kant's theory of emotion on the basis of what Kant says about such evaluations. Our reconstructed theory will not count as a theory of emotion by those who deny that emotions are evaluations of the assumed kind. Moreover, if Kant's view is that the mental states we call "emotions" are *not* evaluations (or not evaluations of the assumed kind), then our reconstruction would not be Kant's theory of emotion; it would rather be a reconstruction of the theory of emotion Kant would have had (or should have had), if he shared our understanding of the nature of the mental states we call "emotions". (a) is therefore problematic, insofar as it would lead to a theory that would not be considered a theory of emotion by some, and may be different than Kant's theory of emotion.

(b) would be a more promising strategy than (a), if there were certain terms that we could all agree designated emotions. For example, if we agree that "fear" and "anger" designate emotions, we can look at what Kant says about fear and anger, and from this we can start reconstructing a more general theory of emotion. But this strategy also has its drawbacks. For if our reconstructed theory would be based only on those terms which we can agree refer to emotions, it would be incomplete and inaccurate to the extent that there are other terms in Kant that refer to emotions, which, had they been taken into account, would make for a more general and accurate theory that applies to all emotions. This puts pressure to introduce more assumptions about which terms designate emotions in ordinary English. But the more assumptions we introduce about ordinary English, the more dubious it becomes that our reconstructed theory would be a theory of

¹¹ We could also combine these two strategies, and start from assumptions both about the nature of emotions and about which terms designate emotions.

emotion. In cases where it is not obvious whether a certain term designates an emotion in English, we could look at what Kant says about the phenomena designated by this term to decide whether it is an emotion. But in order to do that we must fall back on some (contested) conception of emotion, which would allow us to decide whether Kant's remarks about these phenomena suggest or entail that it is an emotion. So, although (b) is *prima facie* more promising than (a), (b) may also result in a theory of emotion that is not Kant's; and thwarting this risk comes at the cost of introducing more assumptions about ordinary English emotional terms, or about the nature of emotions.

Both strategies therefore ultimately depend on contested assumptions: (a) relies on contested assumptions about the nature of emotions, and for (b) to provide a complete and accurate account, it too will ultimately have to rely on contested assumptions about which terms designate emotions in ordinary English and / or about the nature of emotions. That reconstructions of Kant's theory of emotion must rely on such assumptions may seem trivial, given that Kant does not use the term "emotion" or an obvious equivalent, and that theorists disagree about the nature of emotions; and it may also lead one to conclude that all attempts to understand what Kant thought about emotions are doomed to fail. But I do not think either is the case. The methodological observation about reconstructing Kant's theory is not trivial because, as I shall show in the next section, commentators are often not clear about the tacit assumptions of their reconstructions, which lends the impression that they disagree about what Kant says, when they actually disagree about what emotions are, or about which terms designate emotions. And there is still hope for understanding Kantian emotions because, as I shall argue later on (1.6), what we call "emotions" refers to mental states in Kant, some of which he attributes to the faculty of desire, others to the

feeling of pleasure and displeasure, and so we should reconstruct Kant's theory on the basis of what he says about these two kinds of mental states.¹²

1.3. Frierson's Denial of a General Theory

Consider first Frierson's approach to emotions in Kant. According to Frierson, Kant "does not have a general theory of 'emotions'" (Frierson 2014a, 168), and in that respect, he is "in good company amongst at least some contemporary philosophers, who increasingly recognize that what we call 'emotion' is not a unified phenomenon." (ibid., 186, fn. 2). As I read him, Frierson is saying that Kant has no general and unified theory for the variety of mental states we call "emotions", and so, like some contemporary philosophers, Kant would deny that these mental states have a common nature and can be accounted for by one theory. The following passage provides Frierson's vindication of these claims:

In place of a theory of emotions, Kant developed detailed accounts of various mental states, some of which include what we today call emotions. Thus, for instance, Kant discusses anger (*Zorn*), which he classifies as an 'affect' (*Affekt*) (A 7:255) as well as various kinds of hatred (*Haß*) that are 'passions' (*Leidenschaften*) (A 7:270–1). Depending upon its particular form, sexual desire can be an instinct (*Instinkt*), inclination (*Neigung*), or passion (*Leidenschaft*); and love can be a feeling (*Gefühl*), passion (A 7:266), inclination, or can 'lie in the will [*Wille*]' (G 4:399). Some 'emotions', such as wonder (*Bewunderung*), are arguably cognitive states (KU 5:365). For Kant, no univocal concept covers all these cases and each category under which paradigmatic emotions fit¹³ includes other states that are generally *not* considered emotional. Thus 'affects' include not only rage but also shock and laughter (A 7:262). 'Passion' includes not only hatred but also a principled commitment to pursue honor at any price (A 7:270). 'Feeling' includes moral respect and being in love but also the pain of being pricked by a needle. (Frierson 2014a, 168).

¹² My proposed strategy, it must be admitted, also relies on assumptions that may be contested: that emotions are not objective perceptions, and that some of the mental states we call "emotions" are associated by Kant with the faculty of desire, while others are associated by him with the faculty of feeling. But these assumptions are rather minimal and generally accepted. Or so I will argue.

¹³ "Fits" in the original was amended to render the sentence grammatical.

Frierson is saying here that Kant developed accounts of mental states, some of which we today would call “emotions”. His examples of such accounts are Kant's discussions of “anger”, “hatred”, “sexual desire”, “love”, “moral respect”, and “wonder” (call them “prima facie emotion terms”), so Frierson must be assuming that we would call at least some of the mental states designated by these terms “emotions”. He then proceeds to argue that for Kant, “no univocal concept covers all these cases and each category under which paradigmatic emotions fit includes other states that are generally *not* considered emotional”, and so he must also be assuming that there are certain mental states that belong in the same Kantian category as paradigmatic emotions, but we would not call them “emotions”. This argument may be formulated as follows:

(1, implicit) Prima facie emotion terms (“anger”, “wonder”, etc.) - or some of them - designate emotions in ordinary contemporary English; in contrast, other terms (“shock”, “laughter”, “the pain of a needle prick”, and “a principled commitment to pursue honor at any price”) - or some of them - do not designate emotions in ordinary contemporary English;¹⁴

(2) In Kant, there is no single generic term (“no univocal concept”) which all prima facie emotion terms are species of;

(3) In Kant, the generic terms which paradigmatic prima facie emotion terms are species of apply to terms generally not considered as denoting emotions in ordinary English. For example, “shock” and “laughter” are generally not considered as denoting emotions, but like the

¹⁴ Alternatively, Frierson's assumption might be about the nature of emotions (rather than about ordinary language). Read in this way, his premise is that certain phenomena described by Kant (such as anger, hatred) qualify as emotions as we understand them, and others (the pain of needle prick, laughter) do not. This reading seems implausible, however, given that Frierson does not say that he is assuming a certain theory of emotion, and he does not say why on this theory some phenomena - as described by Kant - qualify as emotions, while others do not.

paradigmatic emotion term “anger”, they are species of “affect”; “a principled commitment to pursue honor” is generally not considered as denoting an emotion, but like the paradigmatic emotion term “hatred”, it is a species of “passion”; and “pain in a needle prick” is generally not considered as denoting an emotion, but like “falling in love” and “moral respect”, it is a species of “feeling”.

So, (4) Kant does not have a general theory of emotions.

The argument's first premise, as I have formulated it, is ambiguous, in that it leaves unclear just which terms designate emotions in ordinary English and which ones do not. If Frierson means that all prima facie emotion terms designate emotions, this premise is dubious. For while it is hard to deny that “anger”, “love” and “hatred” designate emotions, it is not obvious that “sexual desire” and “wonder” do. If the premise is merely that some of these terms designate emotions, the argument stands on firmer ground, but then we need to know which ones in order to assess whether from the fact that Kant has no single generic term which these terms are all species of (or from this fact and from 3), it follows that Kant has no general theory of emotion. And a similar problem arises with the claim that some terms do not designate emotions in ordinary English: if the premise is that “shock”, “laughter”, “the pain of a needle prick”, and “a principled commitment to pursue honor”, all do not designate emotions in ordinary English, this premise is dubious. For when we speak of the shock felt by a person who has been assaulted we seem to be referring to an emotion. And although a principled commitment to pursue honor does not seem to qualify as an emotion, this is not an accurate characterization of the passion for vengeance in the passage Frierson is referring to (*Anth.* 7:270). For although Kant says there that

maxims of reason are interwoven with the inclination for vengeance, he does not *identify* passion with a commitment to pursue honor. And since he says earlier that passions *presuppose* maxims (*Anth.* 7:266), it makes better sense to understand the passion for vengeance as *presupposing* a commitment to pursue honor. If this is correct, this passion may be an emotion, and so premise (1) remains dubious (assuming it denies that the term for this passion designates an emotion). The claim in (1) could of course be that only *some* of these terms (“shock”, “laughter”, etc.) do not designate emotions, but then we need to know which ones in order to assess the argument.

Let's assume, though, for the sake of the argument, that Frierson is saying that none of the terms “shock”, “laughter”, “the pain of a needle prick”, and “a principled commitment to pursue honor”, designate an emotion in ordinary English, and that he is right about this. And suppose further, that all these terms are species of generic terms, which paradigmatic *prima facie* emotion terms are also species of. For example, affect is the genus which both laughter and anger are species of. Would this allow us to conclude that Kant has no general theory of emotion? I don't think so. For if, according to Kant, “anger” is an emotion term, and “laughter” is not, and they are both species of the generic term “affect”, this merely means that Kant's theory of emotion is not his theory of affect (for both an emotion term and a term that does not designate emotions are species of the term “affect”).

Frierson does have a point, however, insofar as there should not be too much incongruity between Kant's terminology and ours. For example, we should not identify Kant's equivalent of “emotion” as “pleasure”, if many pleasures in Kant are not emotions in our sense, and if many emotions in our sense are not Kantian pleasures. Some incongruity between Kant's terminology and ours is to be expected, given that more than two centuries stand between us; but the more cases of incongruity we find, the more pressure there is to conclude that we have

misidentified Kant's equivalent of “emotion”. So, although I think that even if true, (1) and (3) should not lead us to deny that Kant has a general theory of emotions, Frierson is right insofar as Kant's equivalent of “emotion” should not apply to too many terms that we would not consider emotional (and insofar as there should not be too many terms that are species of Kant's equivalent term but are not emotional by our standards).

What about the argument's second premise, that Kant has no single generic term which all prima facie emotion terms are species of? The evidence on this issue is actually not as decisive as this premise suggests. For Kant sometimes refers to desire as a “satisfaction” or a “pleasure”, saying, for example, that "Desire is the satisfaction [*Wohlgefallen*] in the existence of the object" (*LA* 25:1334, cf. 15:456); that desire is "satisfaction in the existence of what is within our power" (15:84, my translation); and that "Desire is... a *pleasure insofar as it is a ground of an activity for determining certain representations of the object.*" (*LM* 28:254). Assuming that all mental states of the faculty of desire are desires, these passages are saying that all these mental states are pleasures or satisfactions. And since Kant uses “pleasure” and “satisfaction” as generic terms that refer to all mental states of the faculty of feeling (*Anth.* 7:231), “pleasure” and “satisfaction” might be Kant's generic terms for mental states which all prima facie emotion terms are species of.¹⁵

There is, however, also evidence that desires are causally connected with pleasures or satisfactions, which suggests that they are not pleasures or satisfactions. For example, in a lecture dated 1794-5, Kant says that desiring (*Begehern*) is "*the representation of the object, which is connected with satisfaction in its actuality, and which is the ground of producing it.*" (*LM* 29:1013). And a similar view is expressed in *MM*, where Kant says that "In the case of desire or

¹⁵ I am assuming here that mental states of the cognitive faculty are not emotions. I will argue for this claim in 1.6.

aversion, first, *pleasure* or *displeasure*, susceptibility to which is called *feeling*, is always involved" (*MM* 6:211).¹⁶ According to these passages, pleasure or satisfaction is necessarily connected or involved with desire, but there is no suggestion that desire itself is a pleasure or satisfaction. So, it is not clear whether "pleasure" or "satisfaction" are generic terms which apply to all *prima facie* emotions (be they of the faculty of desire or of the faculty of feeling), or whether they apply only to mental states of the faculty of feeling.

It is important to note, however, that even if such a generic term (or "univocal concept") is missing in Kant, this does not entail that Kant has no general theory of emotion. For it is possible that although there is no concept in Kant that refers to the variety of phenomena we call emotions, Kant nonetheless has some general theory of these phenomena that explains their common nature. The reason he does not explicitly elaborate this theory need not be that this theory does not exist or that Kant denies that emotions have a common nature; it might be that it is simply *of no systemic philosophical interest to him*. In fact, given that Kant's primary systemic philosophical concern is to formulate and vindicate the *a priori* laws of our three mental faculties, it makes sense that a category of emotions (or some German equivalent) would be missing in Kant, if different emotions originate in different faculties (or if certain emotions are the product of more than one faculty), such that there is no *a priori* law that governs all of our emotions.

Before proceeding to look at Cohen's account of Kantian emotions as feelings, it would be good to consider Frierson's reasons not to attribute such accounts to Kant. Frierson warns against the temptation to identify feelings in Kant with emotions, saying: "While distinguishing feeling from cognition and volition is superficially similar to traditional classifications of

¹⁶ I use Grenberg's modified translation (Grenberg 2001, 160, fn. 18). The Cambridge translation is misleading, saying that "...*pleasure* or *displeasure*... is always connected with desire or aversion". But this is just what Kant goes on to deny when he says that there can be a pleasure that is *not* connected with any desire.

‘emotions’ as non-cognitive and volitionally dubious, Kant’s classification of paradigmatic ‘emotions’ under different (often combined) mental faculties belies any easy identification of ‘feeling’ with ‘emotion’. And these intrinsically differentiated states can be related, such that cognitions can cause and provide content for feelings, which can in turn cause volitions."

(Frierson 2014a, 168).

Frierson's point, as I read him, is that it is tempting to think that feelings in Kant are emotions: like emotions, they are non-cognitive (because they are not cognitions) and unreliable as volitions or motivations (because they are not desires or volitions). But the similarity between emotions and feelings in Kant is superficial, for two reasons: first, because paradigmatic “emotions” (i.e., mental states we typically classify as emotions in ordinary language) are classified by Kant as products of different mental faculties (and sometimes more than one); and second, because the mental states produced by the different faculties can be related, both causally and intentionally.

Although Frierson makes this point in arguing against identifying feelings in Kant with emotions, it should be noted that if paradigmatic emotions were classified by Kant as products of more than one faculty, this would also tell against attributing to Kant a single theory of emotion. For Kant takes the three faculties of the mind to be performing essentially different functions, such that there is "always a great difference" between the mental representations they produce (*FI* 20:206); and he speaks of the mental representations of the three faculties as three different kinds of accidents of a single substance, which are irreducible to one another (*LM* 28:145; 29:770). Kant's view thus appears to be that our mental states are of three essentially different kinds, both metaphysically and theoretically, such that cognitions (and intuitions), feelings, and desires have essentially different natures, and are explained by three different theories. And if

this is the case, there is no theory that accounts for mental states produced by the joint effort of the three faculties.

Frierson, however, does not provide evidence that some paradigmatic emotions are classified by Kant as products of more than one faculty. He argues that both fear and sympathetic concern involve a cognitive, a volitional, and a purely subjective feeling component (Frierson 2014a, 172). But from the fact that these mental states *involve* all three components it does not follow that they are *constituted* by all of them, such that they are products of three different mental faculties. And the fact that Kant refers to fear as an affect, i.e., a kind of feeling (*Anth.* 7:255-6), and that he refers to sympathetic joy and sadness at another's state of joy or pain as feelings (*MM* 6:456) means that although they may be closely connected to desires and cognitions, fear and sympathetic concern are products of the feeling of pleasure and displeasure. Of course, these paradigmatic emotions are not isolated events, and their occurrence is normally accompanied by the occurrence of mental states that originate in the other faculties; but this does not entail that they are products of more than one faculty.

Frierson does have a point, however, insofar as some *prima facie* emotion terms designate in Kant mental states associated with the faculty of desire (e.g., hatred, love), while others designate mental states associated with the feeling of pleasure and displeasure (anger, shame). If these mental states form two theoretical kinds, accounted for by two different theories, then Kant has two theories of the mental states we call “emotions”. Granted, a division into two theoretical kinds of emotions is not standard in contemporary theorizing, and is not reflected in ordinary English, but it had been customary among philosophers until roughly the mid 19th century, and some urge that it should be reinstated (Dixon 2012, 339).¹⁷

¹⁷ Kant's desire-feeling distinction does not mirror the old division between the more violent and self-regarding passions, and the milder affections (Dixon 2012, 339), nor does it mirror the division suggested

1.4. Cohen's Account of Emotions as Feelings

Despite Frierson's warnings against identifying feelings in Kant with emotions, Cohen does just that. She shows, however, that she is aware of the difficulty in providing an account of emotions in Kant when she says that her aim is "to extract from Kant's various writings an account of the nature of the emotions and their function - and to do so despite the fact that Kant himself neither uses the term 'emotion' nor offers a systematic treatment of it." (Cohen 2020, 429). Although I find much of what she says about feelings in Kant compelling and well-supported by the text, I think it remains doubtful that it is an account of emotions, for two reasons: (1) it is not clear just which mental states in Kant she thinks are emotions; (2) her vindication of the claim that Kantian emotions are feelings relies on a contested assumption.

Starting with the first problem, it is not immediately clear whether Cohen thinks that only mental states associated by Kant with the faculty of feeling (such as feeling and affect) are emotions, or whether she thinks that mental states associated with the faculty of desire - such as desires, inclinations, passions, and instincts - are emotions too. That mental states of the faculty of desire are *not* emotions on her view is suggested when Cohen says that emotions "originate not from the cognitive faculty or the conative faculty, but from the 'faculty of the feeling of pleasure and displeasure'." (Cohen 2020, 430). Assuming that the mental states that Kant attributes to the faculty of desire originate in the conative faculty, i.e., the faculty of desire, Cohen is excluding such mental states from the domain of emotions (for emotions originate in the faculty of feeling, on her view, not in the conative faculty). One problem with such an exclusion is that some paradigmatic emotions are species of mental states associated with the

by Griffiths, between the more primitive affect programs and the higher cognitive emotions (Griffiths 1997). But the fact that the domain we now recognize as that of emotions included two theoretical kinds until not too long ago, suggests that Kant may have two theories of this domain, without losing their status as theories of emotions.

faculty of desire, which tells against excluding all mental states associated with this faculty from Kant's theory of emotion. For example, hatred is a passion (*Anth.* 7:252; *CPJ* 5:272 fn.), and love, in one of its guises, is an inclination (*GMM* 4:399), in another it is a passion (*Anth.* 7:266). Another problem is that Cohen does not seem to be committed to such a wholesale exclusion because at one point she refers to the inclinations for sex and for freedom as “sensible feelings” (Cohen 2020, 440, fn. 39); and since on Cohen's account “emotion” is co-extensive with “feeling” (*ibid.*, 433, fn. 14), she must think that these inclinations are emotions. In fact, in the following passage, Cohen includes several other mental states that Kant associates with the faculty of desire in the domain of emotions:

In taking emotions to constitute a class of mental states unified by the fact that they are feelings, the interpretation put forward in this paper goes against the recent trend in Kant scholarship that takes the apparent diversity of affective states at face value, emphasizing that what we now call ‘emotion’ includes the numerous, distinct states that Kant calls ‘feeling’, ‘affect’, ‘passion’, ‘inclination’, ‘desire’, and even ‘instinct’. Instead of looking for a single account of the phenomena that we call ‘emotions’, these commentators settle for a mixed model. Yet the problem with this model is that it overlooks the fact that there is something common and unique to emotions, and thereby fails to recognise what makes these states ‘emotions’. In contrast, my aim is to identify what they all share intrinsically. As I will argue, what unifies them in spite of their apparent heterogeneity is the fact that they are essentially ‘feelings’. (Cohen 2020, 432-3)

By “the phenomena that we call ‘emotions’” (lines 6-7) I take Cohen to be referring to the list of Kant's terms “feeling”, “affect”, etc. (call these terms “Kant's affective terms”).¹⁸ If this is

¹⁸ On an alternative reading of the above passage, “the phenomena that we call ‘emotions’” refers not to Kant's affective terms, but rather to some other set of phenomena that we call emotions. Cohen then argues that what unifies these phenomena is that they are feelings. But this suggestion is implausible because Cohen's discussion throughout the paper is focused on *Kant's* affective terms. It might be thought, then, that Cohen is arguing from the premise that Kant's affective terms designate feelings, in some indeterminate sense, independent of Kant's. She then proceeds to reconstruct Kant's theory on the basis of what he says about feelings. But this suggestion is also implausible, for two reasons. First, Cohen says that she will *argue* that what unifies them (presumably, the states designated by Kant's affective

correct, then Cohen is saying that we call the phenomena designated by Kant's affective terms “emotions”.¹⁹ Starting with this assumption, Cohen could then provide an alternative to the trend of “mixed models” (which she associates with Frierson in a footnote). For if the phenomena designated by Kant's affective terms are all emotions, they are all of the same kind. But Cohen also says she will argue that what unifies the affective states designated by Kant's affective terms “in spite of their apparent heterogeneity is the fact that they are essentially ‘feelings’”. And so, she needs to show that all the states designated by Kant's affective terms are essentially feelings. This is certainly true with respect to feeling and affect; but it is doubtful that desire, passion, inclination and instinct, *which Kant associates with the faculty of desire (Anth. 7:251)*, are essentially feelings. And since Cohen's view is that emotions “originate not from the cognitive faculty or the conative faculty, but from the ‘faculty of the feeling of pleasure and displeasure’” (Cohen 2020, 430), saying that desire, passion, inclination and instinct are emotions commits her

terms) is the fact that they are essentially “feelings”, and so she cannot be assuming that Kant's affective terms designate feelings. Second, such an argument would turn on an equivocation of the term “feeling”. For from the fact that emotions are feelings in *some sense*, it does not follow that feelings in *Kant's sense* are emotions. To see why, suppose that emotions are feelings in James' sense (1884), i.e., perceptions of bodily changes. If this is the case, then Cohen should not have reconstructed Kant's theory of emotion on the basis of what Kant says about feelings (for feelings in Kant, on Cohen's understanding, are *not* perceptions of bodily changes). Rather, if James' conception of emotion is correct, Cohen should have reconstructed Kant's theory of emotion on the basis of what Kant says about perceptions of bodily changes. So, the assumption that Kant's affective terms designate feelings *in some indeterminate sense* would not entitle Cohen to reconstruct an account of Kantian emotions on the basis of what he says about feelings. Finally, it might be thought that Cohen does not argue for the claim that Kantian emotions are feelings (rather than some other set of mental states in Kant); she is simply proposing this as one reasonable option among others. But this suggestion conflicts with Cohen's saying that she will *argue* that what unifies these states (the ones designated by Kant's affective terms, or emotions) is the fact that they are feelings; and it also conflicts with her concluding remark “I have *shown* that Kant puts forward a model that defines them (emotions - U.E.) first and foremost as feelings.” (Cohen 2020, 453, italics mine). I therefore think that the reading on which Cohen argues from the premise that we call Kant's affective terms “emotions” is the most plausible.

¹⁹ Alternatively, one might read Cohen as starting from the premise that the mental states designated by Kant's affective terms qualify as emotions, according to a certain theory. This reading seems implausible, however, given that Cohen would then have to first spell out the theory of emotion she is working with, and then show that Kant's affective states qualify as emotions on this theory. Rather than doing that, Cohen proceeds to elaborate Kant's account of feelings.

to the view that these mental states originate in the faculty of feeling. Moreover, since she takes “emotion” to be co-extensive with “feeling” (ibid., 433, fn. 14), Cohen's view must be that only feelings or mental states partly-constituted by feelings are emotions. So, Cohen is also committed to saying that all passions, inclinations, instincts, and desires are feelings or are partly constituted by them.

These two consequences - that certain mental states of the faculty of desire originate in the faculty of feeling and that they are either feelings or are partly constituted by them - might not be as implausible as they seem.²⁰ But even if there is a way to defend Cohen's account despite these consequences, it is important to note that it remains doubtful that her account of feelings is an account of emotions. For Cohen does not argue for the claim that desire, instinct, inclination, and passion originate in the faculty of feeling and that they are either feelings or are partly constituted by them; she is rather assuming that they - along with the other states designated by Kant's affective terms - are all emotions. Our analysis of Frierson's strategy has shown, however, that he denies that all of Kant's affective terms designate emotions, as he denies that some of their species terms - such as “shock”, “laughter”, “the pain of a needle prick”, and “a principled commitment to pursue honor at any price” - are considered emotions. Granted, I have argued above that some of these terms (“shock” and “the passion for vengeance”) may designate

²⁰ Cohen might be willing to accept the idea that certain mental states of the faculty of desire are feelings, as she refers to two mental states of this faculty - the inclination for sex and for freedom - as examples of “sensible feelings” (Cohen 2020, 440, fn. 39). But the idea that mental states of the faculty of desire are feelings (and not desires) seems implausible, while the alternative on which they are *partly-constituted* by feelings has some textual support. For Kant says that desires are causally connected with pleasure or displeasure (*MM* 6:211); and since pleasure and displeasure are feelings, one could argue that desires are partly-constituted by feelings. The claim that the mental states Kant associates with the faculty of desire originate in the faculty of feeling is harder to defend, however. For although it might be argued that sensuous desires and inclinations are caused by pleasure and displeasure and so originate in the faculty of feeling, intellectual or 'sense-free' desires and inclinations *cause* pleasure and displeasure (*MM* 6:212), and so cannot originate in the faculty of feeling. If Cohen takes these desires and inclinations to be emotions too, she is committed to saying that they originate in the faculty of feeling.

emotions after all; but others (“laughter”, and “the pain of a needle prick”) do not seem to designate emotions. If some of these terms do not designate emotions, then Cohen's account is inaccurate, as it includes some affective states we would not call “emotions” (for laughter, pain, and shock, are species of feeling, and so are emotions on Cohen's account).²¹ But more importantly, Cohen's account relies on the assumption that Kant's affective terms designate emotions. And so, her disagreement with Frierson does not turn on textual issues in Kant; it rather stems from the fact that Cohen assumes that we would call all of Kant's affective terms “emotions”, while Frierson denies this much.

1.5. Deimling's "Pragmatic Concept of Emotions"

The last approach to emotions in Kant I wish to discuss is Deimling's. Deimling makes clear she is aware of something similar to our first problem in reconstructing Kant's theory (mentioned at the beginning of 1.2), as she notes that "Kant does not use the term 'emotion' or any other term that picks out what would map onto our contemporary concept 'emotion'." (Deimling 2014, 108). Nevertheless, she argues, "Kant's anthropological writings...provide a concept that is similar in scope. It is implicit in his project of an anthropology 'from a pragmatic point of view' and cuts across different kinds of affective states that Kant distinguishes, such as 'feelings', 'desires', 'affects', 'inclinations' and 'passions.'" (ibid., 108). Deimling calls this concept "Kant's pragmatic concept of emotions".

²¹ Although I do not see Cohen addressing the worry that her account includes mental states that are not emotions, she does admit that it might exclude mental states that are emotions: "it may turn out that some of the states contemporary philosophers routinely class either as emotions or as a necessary part of emotions are not in fact classed by Kant as feelings." (Cohen 2020, 433, fn. 16). If there are indeed such mental states, her account both includes mental states that we would not call “emotions”, and excludes mental states we would call “emotions”.

Which affective states in Kant does Deimling think are instance of his pragmatic concept of emotions? Although the answer is not immediately obvious, one thing that is clear is that some feelings and some desires belong to Kant's pragmatic concept, according to Deimling. This is made clear by the fact that on her view, "feeling" and "desire" are the most basic terms in Kant's taxonomy of affective states, and affect is a kind of feeling, while inclination and passion are desires (Deimling 2014, 109-110). And since her list of affective states which Kant's concept "cuts across" includes both species of feelings ("feeling", "affect") and species of desires ("desires", "inclinations", "passions"), her view must be that at least some feelings (including some affects) and at least some desires (including some inclinations and passions) belong to Kant's concept.

Furthermore, it is clear that *not all* feelings and desires in Kant belong to his concept of emotion. For Deimling takes two criteria to be constitutive of this concept, such that all the affective states that belong to it "(1) track complex values and (2) are under our mediate control." (Deimling 2014, 108). In other words, only those feelings and desires in Kant which track complex values and are under our mediate control are instances of Kant's concept of emotions. And although she does not give a complete list of the feelings and desires in Kant that meet these criteria, she does give us some examples: joy and sadness, a feeling of weariness of life, anger, and compassionate natural feelings. Deimling also claims that Kant's concept "maps well onto our contemporary concept of emotions" (ibid., 121), and so she must think that our contemporary concept of emotions is of those affective states that track complex values and are under our mediate control. I therefore take her strategy for reconstructing Kant's concept of emotions to be starting from the assumption that emotions are affective states that track complex values and are

under our mediate control, and then proceeding to reconstruct Kant's concept of emotions on the basis of what he says about such affective states.

The disagreement on the nature of emotions might lead us to expect that Deimling's account would rely on highly-contested assumptions, but Deimling has found a nice way around the main controversy about the nature of emotions because the assumption that emotions track complex values and are under mediate control seems to be acceptable independently of whether we think emotions are evaluations, feelings, motivations to action, or some combination of these phenomena. For an emotion could track complex value by virtue of consisting in an evaluation, but it could also perform this function by virtue of *being affected by* such evaluations, which are external to it (and so it could be a disposition or a feeling, to the extent that they can be affected by such evaluations). And Deimling's second assumption also seems acceptable on all accounts of emotion, as some degree of mediate control seems possible with respect to feelings, evaluations and motivations. I see, however, two main problems with Deimling's account: (1) her assumptions about the nature of emotions - although uncontroversial - are vague, and once clarified, it might turn out that Kant's pragmatic concept of emotions does not map well onto ours; (2) her evidence does not support the claim that joy and sadness track complex values, which suggests that reconstructing Kant's account on the basis of extra-textual assumptions about the nature of emotions makes it difficult to decide which phenomena in Kant qualify as emotions.

Starting with the latter problem, consider the following passages, which Deimling takes as evidence that joy and sadness track complex value:

All animals are capable of enjoyment [Vergnügen] and pain, but not of joy and sadness; because the latter can only spring from a comparison of the current

situation with our previous situation; but an animal is not capable of making such a comparison. (*LA* 25:422; Deimling's translation 2014, 114)²²

Animals are able to experience neither real joy nor sadness; because these presuppose reflecting on the state after which one is conscious of the current and the preceding state. (*LA* 25:1082; Deimling's translation 2014, 114)

Commenting on these passages, Deimling says: "When we are feeling joy and sadness we are not feeling isolated pleasures and pains, but we are experiencing whatever we are feeling as mattering to a larger part of our lives. We compare our current situation and the feelings associated with it with the 'previous' or 'preceding' situation and the feelings associated with it. And we have a complex feeling in response to considering this context as a whole." (Deimling 2014, 114-15). Deimling therefore takes these passages as evidence that joy and sadness are complex in the sense that they involve comparing our present state and feeling with past states and the feelings associated with them.

This reading might seem to be supported by Kant's claim in the first passage that joy and sadness (the referent of "the latter") can only "spring" from a comparison of the current situation with our previous situation. But note that in the second passage, Kant says that real joy and sadness "presuppose" reflection. This formulation is ambiguous. It could mean, as Deimling reads it, that joy and sadness involve or follow upon reflecting on one's state. But once we consider other pertinent evidence, it becomes clear that joy and sadness cannot follow upon or involve such reflection because they *hinder our ability to reflect*. Joy and sadness are affects (*Anth.* 7:254, *LA* 25:589), according to Kant, and affects hinder our ability for rational reflection: "the feeling of a pleasure or displeasure in the subject's present state that does not let him rise to

²² I have amended Deimling's translation of *Vergnügen* from "pleasure" to "enjoyment", as it is translated in the Cambridge Edition. According to Kant, enjoyment is a species of pleasure (*Lust*): it is a sensuous pleasure through sense (*Anth.* 7:230). This change does not bear on Deimling's argument, however.

reflection (the representation by means of reason as to whether he should give himself up to it or refuse it) is *affect*" (*Anth.* 7:251); "Affects belong to *feeling* insofar as, preceding reflection, it makes this impossible or more difficult." (*MM* 6:407).

If affects make reflection difficult and even impossible, they cannot involve or consist in the reflective activity of comparing one's situation with the past and the future. What Kant means when he says that joy and sadness "spring from" comparison, or that they presuppose reflection, I propose, is that experiencing such feelings, which undermine one's ability to reflect, presupposes the *ability* to reflect. If this is correct, the reason that only rational beings experience joy, sadness and other affects is not that these feelings involve reflection; rather, the reason is that it makes sense to speak of affective states that undermine reflection only with respect to creatures capable of reflection. This is not to deny that tracking complex value is a necessary condition on an affective state's being an emotion, nor that some affective states in Kant track complex value; the point is that the above passages do not show that joy and sadness involve reflection or track complex value, and so Deimling has not provided evidence that joy and sadness are emotions, on her understanding of them.²³ And the broader lesson to be learned here is that reconstructing Kant's theory of emotion on the basis of extra-textual assumptions about the nature of emotions makes it difficult to decide which phenomena in Kant qualify as emotions. Needless to say, any misidentification of such phenomena would distort our account of Kantian emotions.

²³ Deimling's other example of a feeling that tracks complex value - weariness of life - is more convincing. Kant says that he experienced this feeling when he was young because he has a natural disposition to hypochondria due to his narrow chest. He reports that later on he mastered the influence of this condition on his thoughts and actions, "by diverting my attention from this feeling, as if it had nothing to do with me." (*SF* 7:104). Assuming that the feeling of weariness arose and then went away due to a change in Kant's evaluation of the threat to his overall health (rather than due to a physiological change), Deimling is right to take this feeling as tracking complex value.

The other problem I see with Deimling's account is more directly concerned with her strategy of reconstruction. Recall that on Deimling's view, Kant's concept of pragmatic emotions is constituted by two criteria: tracking complex value, and being within our mediate control. If I am reading her correctly, she is assuming that emotions are affective states that track complex value, and are within our mediate control. Accepting Deimling's account of Kant's concept of pragmatic emotions as an account of emotions therefore requires accepting the assumption that emotions track complex value and are within our mediate control. But just how complex is the value that an affective state tracks supposed to be to qualify as an emotion? And what degree of mediate control should it be susceptible to? Although Deimling does not explicitly answer these questions, her examples of value-tracking emotions involve a rather sophisticated reflection process of comparing one's present state with the past or some broader context. But some affective states that we call "emotions" do not involve such a complex process. Consider for example, my fear when being startled by a roaring thunder in the middle of the night. This affective state is sudden and unexpected (assuming I had not seen the preceding flash of lightening), and does not involve comparing my present condition with the past (or future): there is simply not enough time for such a comparison. But my fear seems to be an emotion, and since it is not an emotion by the criteria of Kant's pragmatic concept, there is some discrepancy between this concept and ours. In other words, in her first criterion, Deimling is setting the bar too high, in a way that might unjustly exclude certain affective states in Kant from her account.

Regarding the criterion of being amenable to mediate control, Deimling grants that some affective states, such as moods, are not as amenable to mediate control as our emotions. And so, being amenable to mediate control comes in degrees, on Deimling's view. It is not clear, however, just how much and what kind of control we must be able to exercise over an affective

state for it to qualify as an emotion. Her examples of exercising such control in Kant include changing one's bodily state so as to reduce anger, getting children accustomed to smiling so as to create a cheerful disposition, diverting one's attention from the breaking of a rare goblet in order not to give oneself away to the feeling of pain, controlling our imagination, and bringing certain circumstances to one's attention so as to cultivate compassionate feelings. Should an affective state be amenable to all these kinds of mediate control if it is to qualify as an emotion?

Probably not. Joy, for example, often does not wane when we sit down (in fact, even anger does not always respond to changes in one's bodily state); and grief is awfully resistant to diverting attention from one's loss. Which of these manners of mediate control should an affective state be amenable to, then? Should different affective states be amenable to different modes of control? And at what point does an affective state's resistance to mediate control warrant saying that it is not an emotion?

Until we know the answer to these questions, we do not know what is built into Deimling's assumption that emotions are amenable to mediate control, and so it remains doubtful that Kant's pragmatic concept of emotions maps well onto ours. So, although Deimling offers an interesting alternative to the strategies which start from assumptions about which terms designate emotions in ordinary English, her assumption that emotions track complex value may unjustly exclude some mental states in Kant from her account of emotions; and her assumption that they are amenable to mediate control needs to be spelled out so that we can assess whether Kant's pragmatic concept of emotions is similar to ours.

1.6. Emotions as a Set of Feelings and Desires

To recap our analysis of different approaches to emotions in Kant, we have seen that Cohen thinks that Kant has a general theory of emotion, Frierson denies this much, and Deimling thinks he has a concept that maps well onto our concept of emotion. As for the mental states in Kant that count as emotions, Cohen's view is that they are feelings, but also include mental states that Kant associates with the faculty of desire, such as inclinations and passions; Deimling holds that they are affective states that track complex value and admit of mediate control, and they include both feelings and desires; and Frierson thinks that they may originate in all three mental faculties, and so may include cognitions, desires, and feelings (and combinations thereof). In terms of method, I have argued that Frierson bases his account on assumptions about which terms designate emotions in ordinary English (and which ones do not); Deimling bases hers on assumptions about the nature of emotions; and Cohen starts from the assumption that we would call certain affective states in Kant “emotions”.

Is there hope for providing a cogent and generally accepted account of emotions in Kant, given that commentators disagree not merely about which mental states in Kant qualify as emotions, but also on whether there is a general theory that applies to all these mental states? I think there is. The way to do that, I propose, is to start from the assumption that what we call “emotions” refers in Kant to a set of mental states, some of which he associates with the faculty of desire, others with the faculty of feeling. This assumption is correct, I think, because the mental states associated with the cognitive faculty in Kant - i.e., cognitions in their guise as concepts and intuitions - are objective perceptions (A320/B377), and objective perceptions are not emotions. And since some of the mental states we call “emotions” are subspecies of mental states associated by Kant with the faculty of desire (e.g., love in one of its guises is an

inclination, hatred is a passion), while others are subspecies of mental states associated with the faculty of feeling (anger and shame are affects), some of the mental states we call “emotions” are associated by Kant with the faculty of desire, others with the faculty of feeling.

This proposed strategy relies on one premise about the nature of emotions - namely, that they are not objective perceptions - and on another premise about emotional terms - namely, that at least some of the terms associated by Kant with the faculty of desire (such as love and hatred) and at least some of the terms associated by him with the faculty of feeling (such as anger and shame) designate emotions in ordinary English. But these premises seem to me hard to deny. Moreover, the claim that what we call “emotions” refers to mental states, some of which are associated with the faculty of desire, others with the faculty of feeling, is accepted by Cohen (who treats affective states of both faculties as emotions), Deimling (whose account cuts across different feelings and desires),²⁴ and other prominent commentators.²⁵

Frierson might want to reject this strategy, on account of excluding cognitive states from the domain of emotions. For according to Frierson, "Some ‘emotions’, such as wonder (*Bewunderung*), are arguably cognitive states" (Frierson 2014a, 168). But Frierson's evidence does not support the suggestion that wonder is a cognitive state. In support of his suggestion, Frierson's refers to a passage where Kant says that-

(...) **astonishment** (*Verwunderung*) is a mental shock at the incompatibility of a representation and the rule that is given through it with the principles already

²⁴ Note that this strategy does not presuppose that Kant has a single theory of the mental states we call “emotions”. In fact, for the reasons provided in 1.3, I tend to think that Kant has two theories of these mental states: one for those associated with the faculty of desire, and another for those associated with the faculty of feeling. But my starting premise remains neutral on the issue of whether all the mental states that we call “emotions” - both those associated with the faculty of desire and those associated with the faculty of feeling - have a common nature and can be explained by a single theory.

²⁵ Other commentators who accept this claim include Sherman (who takes Kantian emotions to include feelings, affects, and passions, Sherman 1997, 121 fn. 1); Borges (who takes them to include passions, moral feelings, and affects, Borges 2004, 143); and Sorensen (who take them to include affects, feelings, some inclinations and desires, and some passions, Sorensen 2002, 110).

grounded in the mind, which thus produces a doubt as to whether one has seen or judged correctly; but wonder [*Bewunderung*]²⁶ is an astonishment that continually recurs despite the disappearance of this doubt. The latter is consequently an entirely natural effect of that purposiveness observed in the essence of things (as appearances) (*CPJ* 5:365)

Frierson takes this quote to be evidence that wonder is “arguably” a cognitive state. But it isn’t evident that it is. The doubt, said here to be outlived by wonder, is a cognitive state, but that doesn’t show that wonder is. And there is evidence that wonder is an affect: Kant says earlier that astonishment is “an affect in the representation of novelty that exceeds expectation”, while wonder is “an astonishment that does not cease when the novelty is lost” (*CPJ* 5:272). If wonder is a kind of astonishment, and the latter is a kind of affect, then wonder too is a kind of affect and originates in the faculty of feeling, not in the faculty of cognition. So, Frierson's evidence does not show that some of the mental states we call “emotions” are cognitive states, according to Kant.²⁷

Let's assume, then, that what we call “emotions” refers to mental states in Kant, some of which are associated with the faculty of desire, others with the faculty of feeling. One way to proceed would be to reconstruct Kant's theory of the mental states we call “emotions” by looking at what he says about the different species of the mental states we call “emotions”, associated with these two faculties (i.e., hatred, shame, anger etc.). But as we've seen (1.2), in order for such an account to be complete and accurate, we would have to rely on many assumptions about which terms designate emotions in ordinary language and which terms do not; and the more

²⁶ I have amended the Cambridge translation according to Frierson's for the sake of engaging with his argument.

²⁷ Against my proposed strategy, Frierson might also object that some of the states we call “emotions” are products of more than one faculty in Kant. But I have argued in 1.3 that the evidence does not support this idea, and so I will be assuming here that each emotion is the product of a single mental faculty, that is, of the faculty of desire or the faculty of feeling.

assumptions of this kind we introduce, the more dubious it becomes that our reconstruction is of Kant's theory of emotion. Rather than doing that, I propose that we begin our reconstruction by looking at the two generic terms which all the mental state terms we call "emotions" are species of, namely, "feeling" and "desire". This way, we would get the necessary conditions on a Kantian mental state's being an emotion (qua feeling, and qua desire). Granted, if - as our analysis of Frierson suggests - some of these mental states are *not* emotions (e.g., the pain of a needle prick, laughter),²⁸ then our account would not provide the sufficient conditions on a mental state's being a Kantian emotion. But this is a price we have to be willing to pay if we are to provide an account that rests on minimal and generally accepted assumptions.

²⁸ Frierson's needle prick example is actually a particular case of pain, which is a sensuous displeasure through sense, according to Kant (*Anth.* 7:231). We would probably not call many other sensuous pleasures and displeasures through sense - in food, or in tactile sensations - "emotions". I therefore tend to think that being a desire or a feeling is not a sufficient condition on being a Kantian emotion. But for the reasons mentioned above, I will limit myself to providing the necessary conditions on being a Kantian emotion by looking at the nature of feelings and desires.

Chapter 2. More than A Feeling: Kant's Tripartite Account of Pleasure

2.1. Phenomenological Qualities, Dispositions or Evaluations?

On our starting assumption, some of the mental states we call “emotions” are associated by Kant with the faculty of feeling. We can therefore begin our reconstruction of Kant's theory of the mental states we call “emotions” by looking at what he says about the nature of the mental states or representations associated with this faculty.²⁹ Unfortunately, interpreting Kant on this matter is no easy task, for two reasons. First, because he has several different terms for the genus to which all these mental states belong, and some of them have more than one sense: “feeling” (*Gefühl*), which may also refer to a mental susceptibility or faculty, i.e., to the feeling of pleasure and displeasure;³⁰ “pleasure” (*Lust*) and “satisfaction” (*Wohlgefallen*), both of which sometimes refer solely to “positive” instances of feeling; and arguably “sensation” (*Empfindung*), which also has a narrow sense, different from feeling.

The second reason that understanding the nature of the mental states associated with the faculty of feeling is not easy is that Kant characterizes them in different ways throughout his writings. In some passages he speaks of them as feelings, or as non-cognitive, subjective mental states (e.g., *MM* 6:212, *CPJ* 5:189, 206); in others they are characterized as representations of agreement (pleasure) and disagreement (displeasure) with "subjective life conditions", or with the faculty of desire (*CPrR* 5:9 fn., *LM* 28:894); and in yet other passages, they are explained as grounds for certain activities (*CPJ* 5:220, *FI* 20:230-2).

²⁹ As mentioned in the previous chapter, I am using the terms “mental states” and “mental representations” (or simply “representations”) as roughly equivalent, to denote something present to the mind. In calling the mental states associated with the faculty of feeling *representations* I am therefore not deciding the question of whether they have representational content.

³⁰ As mentioned in the previous chapter, Kant sometimes refers to the feeling of pleasure and displeasure as a mental susceptibility or receptivity rather than as a faculty, to emphasize that it can be passive with respect to its representations.

It is therefore hardly surprising that commentators provide different accounts of feelings or of pleasures in Kant, emphasizing one or more of the aforementioned strands in his writings. In his early work, Guyer defends a phenomenological account on which "all feelings of pleasure are a qualitatively uniform kind of sensation." (Guyer 1997, 105). On this reading, although feelings of pleasure may differ in causal history, duration and intensity, there is "phenomenological uniformity of different feelings of pleasure" (ibid.), such that all pleasures (+) feel the same,³¹ and all displeasures feel the same.

Other commentators provide accounts of *feelings* in Kant, but they are meant to apply to the same phenomena as Guyer's, that is, to the mental states Kant associates with the feeling of pleasure and displeasure.³² DeWitt, for instance, argues that feelings in Kant are "action-initiating evaluative judgments" (DeWitt 2018, 74), and so her account may be characterized as an evaluative-dispositional account. Cohen also stresses that feelings are evaluative, but on her account, the other essential component of feelings is not dispositional or motivational, but rather phenomenological. According to Cohen, feelings are "affective appraisals of our activity" (Cohen 2020, 430), and different types of feelings have different phenomenological characters, irreducible to differences in intensity (ibid., 436).³³

³¹ Recall that a plus sign indicates that the term preceded by it refers solely to "positive" mental states.

³² Cohen's account (2020) is probably meant to apply to mental states associated with the faculty of feeling, and to some mental states associated with the faculty of desire. For, as I have argued in the previous chapter, she takes certain mental states associated with the faculty of desire (i.e., inclinations, passions, instincts, and desires) to be emotions. Since Cohen also thinks "feeling" is co-extensive with "emotion", her account of feeling is probably meant to apply to some mental states associated with the faculty of desire.

³³ For her earlier and slightly different formulation of this account, see Cohen 2017, 678-80. Although both DeWitt and Cohen take feelings to be evaluative, for DeWitt the evaluation of feelings is done by means of judgment, while Cohen argues that "rather than conveying their appraisals through representational content, as judgment does, they convey it through their valence - either pleasurable or painful." (Cohen 2020, 438). For two other accounts, similar to Cohen's in emphasizing the difference between feelings and cognition, see Deimling 2018, and Grenberg 2001.

Finally, in his more recent work, Guyer argues that in some passages, Kant speaks of pleasure as a behavioral disposition. He still insists that "Kant often writes of pleasure and displeasure or pain³⁴ as two distinctive types of feeling... qualitatively or phenomenologically identical across all of their own tokens or instances" (Guyer 2018, 147); but Guyer now adds that Kant sometimes writes "as if pleasure and pain are not feelings at all, and therefore not a single way that all pleasures feel and not a single way that all pains feel, but rather as if pleasure and pain are more like behavioral dispositions than distinctive sensations." (ibid., 148). On this "dispositional model", pleasure is a disposition to remain in a state one finds pleasant, and displeasure is a disposition to exit a state one finds unpleasant (ibid., 149).³⁵

What is the nature of the mental states associated with the faculty of feeling, then? Are they best understood as behavioral dispositions, felt phenomenological qualities, or evaluations? The present chapter argues that Kant's account combines all three components, but in forms that slightly differ from those we find in its aforementioned interpretations. These mental states, I will argue, are representations of objects' agreement and disagreement with our ability or

³⁴ Guyer is using "pain" interchangeably with "displeasure", as the opposite of pleasure. This is common practice, and as Katz notes, no alternative has enduringly replaced "pain" as the inclusive English term opposed to "pleasure" (Katz 2016, fn. 1). On Kant's terminology, however, pain (*Schmerz*) is a species of displeasure: it is sensuous displeasure through sense (*Anth.* 7:230). I nonetheless take Guyer to be providing an account that is meant to apply to all mental states of the feeling of pleasure and displeasure, and not merely to their sensuous species, pain and enjoyment (*Vergnügen*, translated as "gratification" in the Guyer and Matthews translation of *CPJ*). That Guyer's account is not restricted to sensuous pleasures is made clear by the fact that he uses it to interpret Kant's remarks about the experiences of the beautiful and the sublime, which involve non-sensuous pleasures.

³⁵ Similarly, Rueger argues that Kant has an "attitudinal" theory of pleasure, on which "a pleasurable representation is one towards which we adopt an attitude that is variously described as preference or desire." (Rueger 2020, 136). Other accounts are not as easily classified. Frierson, for instance, attributes to Kant a dispositional account, but also hints at the phenomenological and evaluative aspects of pleasure: "when one feels pleasure, one feels like continuing in one's state because one's state seems conducive to the activity of one's powers." (Frierson 2014b, 58). Newton's seems to be a hybrid evaluative-phenomenological account, as she says that "pleasure in humans is a feeling of agreement with open-ended and indeterminate capacity to determine our own form of life through reason, broadly understood as a capacity for cognition." (Newton 2017, 526).

intention to act, that have a felt quality, and dispose us to remain in our state or exit it, and to act in order to produce or not to produce their objects.

I start with some background on the faculty of feeling and the different terms Kant uses for the genus to which all its mental states belong (2.2). I proceed to look at Kant's comments on the subjective, non-cognitive nature of pleasure, and argue that they should be read as saying that pleasure involves a representation of an object's relation to the subject, and a felt quality (2.3). There is also evidence, however, that pleasure involves a behavioral disposition, and I propose that Kant's shifting between these different characterizations is best interpreted by attributing to him a tripartite account. Such an account is both supported by the text and philosophically attractive because it can explain how the three functions traditionally associated with pleasure all hang together in one mental state (2.4). I end by comparing my account with Guyer's and Cohen's (2.5), and considering some objections to it (2.6).

2.2. Pleasure, Satisfaction, and Feeling

On Kant's mature picture of the human mind, there are three primary mental faculties: the faculty of cognition, the faculty of desire, and the feeling of pleasure and displeasure (*FI* 20:245; *CPJ* 5:177; *Anth.* 7:123-4). My aim here is to provide an account that applies to all the mental states associated with the faculty of feeling. But in order to do that we must first get clear on Kant's generic terms for the mental states of this faculty, which, I shall argue, include “pleasure” and “displeasure”, “satisfaction” and “dissatisfaction”, and “feelings”.

That the mental states associated with the faculty of "the feeling of pleasure and displeasure" are all species of pleasure or displeasure is not only suggested by this faculty's

name, but also made clear by Kant's opening paragraph of Books II of his *Anth.*, which provides the following "Division" of the feeling of pleasure and displeasure:

1) *Sensuous pleasure*, 2) *intellectual pleasure*. The *former* is either introduced A) through *sense* (enjoyment), or B) through the *power of imagination* (taste); the second (that is, intellectual pleasure) is either introduced a) through representable *concepts* or b) through *ideas*, - - and thus the opposite, *displeasure*, is also introduced in the same way. (*Anth.* 7:230)

This passage provides an exhaustive taxonomy of the types of the mental states or representations produced by the feeling of pleasure and displeasure. On this taxonomy, this faculty's representations are either sensuous (*sinnlich*, also translated as "sensible" or "sensual") or intellectual, depending on the way they are "introduced", which I take to mean that they arise in response to representations of different mental faculties. Some of these mental faculties, namely, the understanding and reason, are intellectual, and so are the pleasures (+) in response to their representations;³⁶ other mental faculties (sense, the power of imagination) are sensuous, and so are the pleasures (+) in response to their representations. And since Kant says that displeasure is "introduced in the same way", it is reasonable to infer that displeasure may also be sensuous or intellectual. So, Kant is providing here the following exhaustive taxonomy of the mental representations of the feeling of pleasure and displeasure: sensuous pleasure (+), intellectual pleasure (+), sensuous displeasure, and intellectual displeasure. All types and tokens of the

³⁶ Frierson argues that Kant generally "treats the terms '(dis)satisfaction' and '(dis)pleasure' as synonyms, but he distinguishes them in the context of discussing moral motivation." (Frierson 2014b, 57, fn. 13). His view seems to be that in the context of moral motivation, "pleasure" and "displeasure" designate sensuous species of satisfaction and dissatisfaction, for he says that when caused by sensibility, satisfaction and dissatisfaction "are properly called pleasures...and displeasures" (Frierson 2014b, 160). While I think some of Frierson's evidence does not support this claim, there is no need to argue for this here, as I will not be basing my account of the mental states of the faculty of feeling on what Kant says about pleasure and displeasure in the context of moral motivation.

mental representations of the faculty of feeling are therefore species (and subspecies) of pleasure or displeasure.³⁷

Kant uses another pair of generic terms, however, to designate all “positive” and “negative” mental states of the faculty of feeling, be they intellectual or sensuous, namely, “satisfaction” (*Wohlgefallen*) and “dissatisfaction” (*Mißfallen*). This is made clear in the following passage:

-- The expressions for what *satisfies* and what *dissatisfies* [*was gefällt oder mißfällt*] and for what is in between, the *indifferent*, are too *broad*; for they can also refer to intellectual ones [*aufs Intellectuelle gehen*]; where they would then not coincide with enjoyment and pain. (*Anth.* 7:230)

Although he uses them in their conjugated verb forms, Kant's point here is that the terms “satisfaction” and “dissatisfaction” are not equivalent with “enjoyment” and “pain” because the latter are sensuous pleasure and displeasure (respectively), while the former “can also refer to intellectual ones”, meaning that they could be either intellectual or sensuous. “Satisfaction” and “dissatisfaction” are therefore Kant's generic terms for the mental states of the faculty of feeling, such that all the mental states of this faculty are species of satisfaction or dissatisfaction. And that “satisfaction” and “dissatisfaction” are generic terms for mental states of the faculty of feeling is made clear by the fact that Kant calls this faculty “the faculty of satisfaction and dissatisfaction” (*LM* 29:890; or the faculty “of being well-pleased and displeased”, on the Cambridge translation of *LR* 28:1059).

³⁷ When I say that all types and tokens of the mental representations of the faculty of feeling are species of pleasure or displeasure, the “or” should be taken in an inclusive sense. For Kant's view is that respect for the moral law and the feeling of the sublime are combinations of pleasure and displeasure (*CPrR* 5:77-8, *CPJ* 5:260). If “respect” and “the feeling of the sublime” refer to individual mental states of the feeling of pleasure and displeasure (rather than to two distinct states of this faculty), then some of the mental states of this faculty are a combination of pleasure and displeasure. Moreover, Kant's view appears to be that some of the mental states of this faculty are pleasures which are partially intellectual, partially sensuous (this is how he characterizes the feeling for the beautiful in *Anth.* 7:239.)

Another term Kant uses to refer to mental states of pleasure and displeasure is “feeling”. That “feeling” can refer to such mental states is suggested by the fact that pleasure and displeasure are produced by the faculty of *feeling*; and this impression is reinforced by the fact that Kant refers to different kinds of pleasure and displeasure as feelings. For instance, he refers to "sensuous pleasure in the sensation of an object" as "the feeling for the agreeable" (*Anth.* 7:230) and to pleasure in reflective intuition as "the feeling for the beautiful" (7:239); he calls sensuous pleasure introduced through sense “enjoyment” (*Vergnügen*, "gratification", in the Cambridge Edition of *CPJ*), and says that it is "the feeling of the promotion of life" (7:231); pleasure (as well as enjoyment) is said to be "The feeling of the promotion of life", displeasure "the feeling of the hindrance of life" (*LA* 25:559; cf. *LM* 24:247). Moreover, Kant says that "we will call that which must always remain merely subjective and absolutely cannot constitute a representation of an object by the otherwise customary name of 'feeling'" (*CPJ* 5:206), and “feeling” here is naturally read as referring to a mental representation or state (rather than to a mental faculty); finally, Kant characterizes “feeling” as a state of pleasure or displeasure when he says that "Feeling is the state where one feels pleasure or displeasure." (*LA* 25:1340). These passages make clear that “feeling” in Kant can refer to types and tokens of mental states associated with the faculty of feeling. But Kant also uses “feeling” to refer to a mental faculty (for example, when he says that "The capacity for having pleasure or displeasure in a representation is called *feeling*", *MM* 6:211; cf. *FI* 20:117). And so, in every passage where this term comes up we should try and determine whether it refers to the faculty of feeling or to a mental state produced by it.³⁸

³⁸ Kant warns against calling the faculty of pleasure and displeasure “the feeling of pleasure and displeasure”, and proposes that it is better called “the faculty of satisfaction and dissatisfaction”, "since the word *feeling* appears to connote something sensible" (*LR* 28:1059). But although he says “feeling” *appears* to connote something sensible, he uses it as a generic term that refers to both sensible and

Feelings, in their guise as mental states of pleasures and displeasures, have many species and subspecies. For in addition to his distinction between sensuous and intellectual pleasures, Kant attributes a wide variety of mental states to the faculty of feeling, including mental movements or agitations of the mind (*Gemüthsbewegung*),³⁹ emotions (*Rührung*),⁴⁰ affects, and special types of affects, such as fear, courage, anger, shame, laughing and crying (*Anth.* 7:256-264). But these are all species and subspecies of the mental states of the feeling of pleasure and displeasure, and so the account I will now develop is meant to apply to all of them.⁴¹

intellectual mental states, and to refer to a mental faculty (which produces mental states of both kinds). Another name Kant uses for this faculty is “interior sense” (*Anth.* 7:153; *LM* 29:890).

³⁹ Thomason bases her account of Kantian emotions on what Kant says about *Gemüthsbewegung* (Thomason 2017, 450-7). Kant, however, characterizes *Gemüthsbewegung* as “a feeling of pleasure or displeasure that draws our entire attentiveness to it.” (*LA* 25:1340); and he says that *Gemüthsbewegung* is a state “where the mind does not have its sensations and desires in its control.” (*LA* 25:589), and proceeds to distinguish between two kinds of *Gemüthsbewegung*: passions, and affects. An account of emotions based solely on *Gemüthsbewegung* is therefore inaccurate because some emotions do not draw our entire attentiveness to them, and do not involve having no control over our sensations and desires. Thomason addresses the worry that her account applies only to affects, but the reason she presents for this worry is that a passage she cites as evidence for her account appears in the section of *Anth.* on affects. My reason to think that her account applies only to affects and passions is different, however, namely, that Kant’s characterizations of *Gemüthsbewegung* do not hold for all emotions.

⁴⁰ *Rührung* is explained by Kant as “a sensation in which agreeableness is produced only by means of a momentary inhibition followed by a stronger outpouring of the vital force” (*CPJ* 5:226). This term is translated in the Cambridge edition as “emotion”, but this is misleading because in contemporary English “emotion” does not have the specialized sense of *Rührung*, and refers to a broader range of affective states (see Sorensen 2002, 110, for a similar point). Indeed, the fact that Kant speaks of *Rührung* as a sensation in which agreeableness is produced suggests that it is a *sensuous* pleasure. And since not all emotions are sensuous pleasures, *Rührung* should not be translated as “emotion”.

⁴¹ Kant also allows that we may call a determination of the feeling of pleasure and displeasure “sensation” (*Empfindung*), as long as we keep in mind that in contrast to a sensation through sense, which is related to the object, sensations in their guise as pleasures are “related solely to the subject” (*CPJ* 5:206); and in *FI*, he says that “there is only one so-called sensation that can never become a concept of an object, and this is the feeling of pleasure and displeasure.” (*FI* 20:224). Zuckert argues that Kant does not think that pleasure is a sensation because pleasure has no objective reference (Zuckert 2002, 240). I agree that pleasure has no objective reference, but think there is a sense of “sensation” on which it has no objective reference, and feelings / pleasure and displeasure are sensations in this sense. Rueger denies that pleasures are sensations on the basis of *CPJ* 5:205-7, which he takes to be a *reductio* on the premise that pleasures are sensations (Rueger 2020, 126-7). But since Kant goes on to speak of determinations of the faculty of feeling as sensations which are “related solely to the subject” (*CPJ* 5:206), this passage is better understood as a *reductio* on the premise that all feelings are pleasures *in* sensations, i.e., enjoyments (or “pleasures in the agreeable”). And if this is correct, *CPJ* 5:206 is no evidence that pleasures or feelings are not sensations. Finally, according to DeWitt, “Kant uses the term *Empfindung*, as opposed to the more

2.3. Pleasure as a Subjective Representation and a Felt Quality

Since all the mental states associated with the feeling of pleasure and displeasure are feelings, and are species of pleasure or displeasure (or of satisfaction and dissatisfaction), we can proceed to reconstruct Kant's account of these mental states on the basis of what he says about pleasure and displeasure, satisfaction and dissatisfaction, and feelings; and for the sake of brevity, I will call this account "Kant's account of pleasure".⁴² The first characterization of pleasure I wish to consider takes it as a subjective representation or as the subjective aspect of a representation:

The capacity for having pleasure or displeasure in a representation is called *feeling* because both of these involve what is *merely subjective* in the relation of our representation and contain no relation at all to an object for possible cognition of it (or even cognition of our condition). While even sensations, apart from the quality (of e.g. red, sweet, and so forth)⁴³ they have because of the nature of the subject, are still referred to an object as elements in our cognition of it, pleasure or displeasure (in what is red or sweet) express nothing at all in the object but simply a relation to the subject. (*MM* 6:211-2)

general term *Gefühl*, in order to refer specifically to sensations of gratification / enjoyment and pain." (DeWitt 2014, 44). But Kant says that the feeling for the agreeable, i.e., enjoyment, is pleasure *in* the sensation of an object (*Anth.* 7:230). So, his view is not that *Empfindung* is a sensuous feeling, such that it refers to enjoyment and pain; it is rather that enjoyment and pain are sensuous pleasure and displeasure, in response to objective sensations.

⁴² I will not base my account on Kant's comments about sensations because there is a sense of "sensation" which is different from "feeling" (see previous footnote), and it is not always clear which sense "sensation" has in a given passage. And while I think that pleasures are subjective sensations, I will not base my argument for the phenomenological component of pleasure on this.

⁴³ Kant seems to have misplaced these parentheses, for their location suggests that color depends on the nature of the subject and so it is a subjective sensation which cannot contribute to cognition. But elsewhere, Kant says that "The green color of the meadows belongs to **objective** sensation, as perception of an object of sense" (*CPJ* 5:206). We can reconcile this apparent contradiction by reading *MM* 6:211-2 as saying that red and sweet are examples of sensations, not of subjective sensations. If this is what Kant meant to say, he should have placed the parentheses after the word "sensations".

In the first sentence, Kant is saying that the capacity for taking pleasure and displeasure in a representation is called “feeling” because these mental states involve what is merely subjective in the relation of our representation and contain no relation that would allow us to cognize an object or our own condition. I take this to mean that the mental states of the faculty of feeling cannot be grounds for cognition: neither cognition of objective properties (that the apple has a certain shape, for example), nor of our own condition as affected by the object (that the apple raises my blood sugar level).⁴⁴ The reason that these mental states cannot be grounds for cognition, is given in the second sentence: it is because they “express nothing at all in the object but simply a relation to the subject”. In other words, pleasure and displeasure do not refer to the objects that elicit them - they rather refer to the subject as it is affected by the object or the representation - and so the judgments based on them fall short of cognition, which on Kant's view, has objective grounds and reference (A320/B377). As Kant explains in the footnote appended to this passage, feeling is "the effect of a representation... upon a subject" that cannot become an element in our cognition because it involves only a relation of the representation to the subject (*MM* 6:211 fn.).⁴⁵

⁴⁴ Cohen refers to Kant's claim that "pleasure or displeasure (in what is red or sweet) express nothing at all in the object but simply a relation to the subject", as an account of *sensible* feelings (Cohen 2020, 440, fn. 39). This reading might appear to be supported by the fact that Kant's examples are of sensible feelings, in sensations of red and sweet. But these are merely *examples* of pleasure and displeasure, and need not be taken to suggest that all pleasure and displeasure are of this sort, nor that what Kant says in this passage applies solely to sensible pleasure and displeasure. Moreover, Kant goes on to speak of contemplative pleasure that is not necessarily connected with desire, and of an intellectual pleasure (*MM* 6:212), which makes clear that not all pleasure and displeasure are sensible; and we've seen that in *Anth.* 7:230, he also speaks of intellectual pleasures. It might be thought that the claim that pleasure and displeasure express simply a relation to the subject does not apply to pure aesthetic and moral feelings, because such feelings ground objective cognitions. But in fact, neither of these feelings grounds objective cognition. Pure aesthetic pleasures ground judgments that make a claim to universal validity, according to Kant (*CPJ* 5:211-217). And moral feeling, although grounded in a judgment of reason (*MM* 6:387), is "something merely subjective, which yields no cognition." (*MM* 6:400).

⁴⁵ In this way, pleasure differs from the subjective formal and material elements of our representations, that is, from pure intuitions and sensations, respectively: pleasure is a subjective element that cannot contribute to cognition, while pure intuitions and sensations are subjective elements that do contribute to

Part of what Kant means in speaking of the subjective, non-cognitive nature of pleasure, I propose, is that pleasure has a *felt phenomenological character*. This felt character is suggested by the fact that pleasure and displeasure are *feelings*, and in passages where Kant speaks of the difficulty in explaining them. For example, Kant says that by means of the relation to the feeling of pleasure and displeasure "nothing at all in the object is designated, but... the subject feels itself as it is affected by the representation." (*CPJ* 5:204). If pleasure and displeasure are states in which the subject *feels itself* as it is affected by the representation, they should have a felt character or quality. Moreover, Kant argues that because pleasure and displeasure are not "kinds of cognition", they "cannot be explained by themselves at all, and are felt, not understood" (*FI* 20:232). The claim that we can feel pleasure and displeasure but not understand them suggests that these mental states have a phenomenological character that does not admit of understanding by means of concepts.⁴⁶ And a similar claim appears when Kant says that pleasure and displeasure "express nothing at all in the object but simply a relation to the subject", and then proceeds to argue that for this reason they "cannot be explained (*erklärt*) more clearly in themselves" (*MM* 6:212). Here, too, the limitation on explaining pleasure and displeasure may be understood as due to their felt qualities. It is this felt quality of pleasure that renders it essentially different from representations of the cognitive faculty, and requires assigning it to a different faculty - the faculty of feeling.⁴⁷

cognition. Although this passage contrasts pleasure with sensation, which suggests that all sensations have objective reference, in *CPJ* 5:206 Kant is allowing that determinations of the faculty of feeling be called "sensations" (see fn. 41). If these passages are to be read as compatible, we had better read *MM* 6:211-12 as contrasting pleasure with *objective* sensation. The reason that pleasures do not contribute to cognition and have no objective reference, I propose, is that they depend on the nature of a *particular* subject; pure intuitions and objective sensations, in contrast, depend on the nature of the transcendental subject, i.e., the subject of experience.

⁴⁶ Deimling also takes *FI* 20:232 as pointing to the phenomenology of pleasure (Deimling 2018, 34).

⁴⁷ Cohen cites other passages as evidence that feelings have a phenomenological character, but some of them need not be read as referring to their phenomenology. According to one such passage, the feeling of

It might be thought that if we cannot explain pleasures by themselves, then they must have no representational content: they are opaque mental states that are not “about” anything, and this is why we struggle to make sense of them. This suggestion might seem to gain support from passages where Kant speaks of pleasure as an aspect of a representation with which it is connected. For example, Kant says that “the subjective aspect in a representation **which cannot become an element of cognition at all** is the **pleasure** or **displeasure** connected with it” (*CPJ* 5:189); and he characterizes sensibility, of which the faculty of feeling is a species, as “the subjective aspect of our representations in general” (*MM* 6:212 fn.). If pleasure and displeasure are *aspects* of representations with which they are connected, it might be thought that they are not representations in their own right, and so they have no representational content independent of their associated representations. On this view, my pleasure in the taste of cold ice cream on a hot summer's day is a felt quality attached to the sensory representation of the ice cream's taste; but my pleasure itself does not represent anything.⁴⁸

respect is “consciousness of the *subordination* of my will to a law” (as cited in Cohen 2020, 436, fn. 27; probably referring to “The consciousness of a *free* submission of the will to the law”, *CPrR* 5:80), which could be understood such that in respect we simply represent our will as subordinate to the will (without this representation having a phenomenological character). According to another passage, the sublime involves “as its characteristic mark a **movement** of the mind” (*CPJ* 5:247), but a movement of the mind is explained elsewhere as “a feeling of pleasure or displeasure that draws our entire attentiveness to it.” (*LA* 25:1340). In saying that the sublime involves a “movement of the mind” Kant is not characterizing its phenomenology, but rather pointing out that it draws our entire attention to it. Finally, Cohen says that Kant approves of Burke’s psychological descriptions of the difference between the phenomenology of the feeling of the beautiful and that of the sublime (*CPJ* 5:277). But although Kant says that Burke's remarks are “extremely fine” as empirical psychological remarks, he contrasts this approach with his own transcendental approach, and so it is not clear he endorses the phenomenological aspects of Burke's account.

⁴⁸ For a defense of the claim that hedonic tone - the quality of the experience of pleasures and pains - only exists as an aspect of otherwise non-hedonic elements of consciousness, see Duncker 1941, 399-400.

In other passages, however, Kant speaks of pleasure as a representation *of* something, and of feeling as drawing certain distinctions. For instance, in a footnote in the preface to *CPrR*, Kant provides the following explanation (*Erklärung*) of pleasure:⁴⁹

Pleasure is the *representation of the agreement of an object or of an action with the subjective conditions of life*, i.e., with the faculty of the *causality of a representation with respect to the reality of its object* (or with respect to the determination of the powers of the subject to action in order to produce the object). (*CPrR* 5:9 fn.)

Kant is saying here that pleasure (+) is the representation of the agreement of an object or an action with the subjective conditions of life, that is, with the faculty of the causality of a representation with respect to the reality of the represented object, or with determination of the subject's powers to act in order to produce the object. I take this to mean that pleasure (+) is a representation of an object as agreeing with one's ability to act according to representations and bring them about, or with one's intention or disposition to do so. The object in question should probably be understood in a broad sense, which may include physical objects (apples, cars), people and non-rational animals, states of affairs (world peace), and actions (jogging, performing one's moral duty). And we may reasonably infer that displeasure is a representation of an object (in this broad sense) as disagreeing with one's ability to act, or with one's intention or do so. But what's important for our present concern is that a representation of an object as agreeing or disagreeing with one's ability or intention to act is not a mere felt quality caused by such agreement or disagreement; it is rather an evaluation of the object's relation to the subject.⁵⁰

⁴⁹ I will address the question of how to reconcile Kant's explanations of pleasure with his restriction on such explanations in 2.6.

⁵⁰ It might be objected that this characterization does not apply to pure aesthetic pleasure (+), which occurs when the understanding is in free-play with the imagination. But the claim that pleasure (+) is a representation of agreement with one's ability or intention to act is about pleasure's representational

That pleasure involves evaluation is also suggested in passages where Kant speaks of feeling as drawing distinctions, and of pleasure as involving judgment. Kant says, for example, that "Feeling makes distinctions only for oneself; the judgment is not valid for others" (*R.* 1850 16:137), indicating that the faculty of feeling (or its produced states) draws certain distinctions, and suggesting that it either makes or grounds judgments that are not valid for others. Similarly, he says that the faculty of pleasure and displeasure is "The faculty of the discrimination of representations, insofar as they modify the subject" (*LM* 28:586), which means that the faculty of feeling distinguishes between objects in accordance with the way they affect the subject. Kant even goes as far as characterizing taste as "a faculty for judging through satisfaction or dissatisfaction" (*LM* 28:249), and the higher faculty of pleasure and displeasure as "the faculty for judging of an object whether it pleases or displeases from cognition of the understanding according to universally valid principles." (*ibid.*). Although these claims suggest that pleasure not only grounds judgments but also consists in a judgment, the judgment in question cannot be a judgment in Kant's usual technical sense. For a judgment in this latter sense is a cognition of an object by means of a concept (*A*68/*B*93, *B*141), and as we have seen, Kant denies that feelings contribute to cognition (*MM* 6:212) and he says that they are not kinds of cognition (*FI* 20:232). But the fact that feeling "make distinctions for oneself" and discriminates between objects on the basis of their subjective affect makes clear that its mental states are subjective evaluations that distinguish between different objects in accordance with their impact on us.⁵¹

content, not about its cause; it is therefore compatible with pleasure being caused by the free-play of the faculties.

⁵¹ One way to understand the kind of evaluation involved in pleasure, suggested by DeWitt (2014, 39), is as a *practical* cognition, i.e., cognition having to do with determining grounds of the will. If this is correct, then in passages where Kant says that pleasures are not cognitions, he means either that they are not theoretical cognitions (that is, they are prescriptive, not descriptive) or that they are not speculative cognitions (such that we can derive imperatives from them) (9:86).

Kant's characterizations of the kind of representations that pleasure and displeasure involve slightly differ: some are more naturally read as referring to a felt quality, others to representational content. For example, in one lecture Kant says that "if a representation harmonizes with the entire power of the mind, with the principle of life, then this is *pleasure*." (*LM* 28:247). Here Kant does not say that pleasure is a representation *of* harmony, but rather that it harmonizes with the entire power of the mind, explicated as the principle of life, which suggests that pleasure (+) has a harmonious felt quality. In contrast, in another lecture he says that pleasure is "the representation of the agreement of an object with the productive power of the soul" (*LM* 29:894), and that pleasure is "the consciousness of the agreement of an object with the productive power of imagination of our soul." (*LM* 29:891).⁵² These last two passages are more naturally read as referring to the content that pleasure (+) represents (agreement with the productive power of the soul, and agreement with the productive power of imagination), than to its felt quality. But rather than choosing between these readings, we may read Kant in a way that combines them, such that pleasure involves both a subjective experience of harmony, and a representation of harmony, i.e., of agreement between an object and one's ability or intention to act. And this reading is plausible, given that the unique representational content of pleasure can explain why it must have a felt character: it is *because* pleasure represents the way the subject is affected or modified, that it cannot be a "cold" objective representation - it rather must have a felt subjective quality.

Before proceeding to look at passages where pleasure is characterized in dispositional terms, I want to consider a possible objection to the inclusion of the phenomenological

⁵² Kant's characterizations also slightly differ in terms of the subjective capacity with which the objects we take pleasure (+) in agrees, but they are all related to action or productivity. This is also the case with the productive power of imagination (*LM* 29:891), understood as a capacity to represent future objects and states of affairs we desire to bring about in action.

component in my account of pleasure. One might object that although pleasure *could* have a felt quality, it need not, and so a felt quality is not essential to it, and should be excluded from its account. DeWitt argues along these lines, saying that if pleasure can be unconscious, "then an unconscious pleasure is a pleasure, not in virtue of its consciously felt qualities... but in virtue of its representational content." (DeWitt 2018, 74, fn. 7). The passage DeWitt cites in support of the claim that pleasure may be unconscious is this:

The feeling that urges the subject to **remain** in the state he is in is *agreeable*; but the one that urges him to **leave** it is *disagreeable*. Combined with consciousness, the former is called *enjoyment (voluptas)*, the latter *lack of enjoyment (taedium)*. (*Anth.* 7:254).

Although this passage does suggest that the feelings of the agreeable and the disagreeable need not be combined with consciousness, elsewhere Kant's view appears to be that "agreeable" is an attribute of objects or representations we find enjoyable (or gratifying, on the Cambridge translation) - not a mental state: "**Agreeable** is that which everyone calls what **gratifies** him" (*CPJ* 5: 209-10). Moreover, according to some passages, pleasure is a *conscious* mental state: "Pleasure is...the consciousness of the agreement of an object with the productive power of imagination of our soul" (*LM* 29:891); "The consciousness of the causality of a representation with respect to the state of the subject, **for maintaining** it in that state, can here designate in general what is called pleasure" (*CPJ* 5:220). If pleasure is a conscious mental state, there is no reason to think that a felt phenomenological character is inessential to it and should be excluded from its account because pleasure and its felt quality cannot come apart.⁵³

⁵³ I therefore propose that Kant's point in *Anth.* 7:254 is not that feelings of enjoyment and lack of enjoyment may be unconscious, but rather either that (1) the urges they involve (to remain in one's state or to change it) may be unconscious; or (2) the properties associated with them (agreeableness / disagreeableness) may be unconscious.

2.4. Pleasure as a Disposition

On my proposed account, pleasure involves both an evaluation and a felt quality. There are also passages, however, where Kant speaks of pleasure and displeasure as grounds for certain activities. For example, Kant provides the following “transcendental explanation” of pleasure, which abstracts from empirical data, and is meant to hold for all pleasures, whether they are accompanied by sensation, reflection, or a determination of the will:

Pleasure is a **state** of the mind in which a representation is in agreement with itself, as a ground, either merely for preserving this state itself (for the state of the powers of the mind reciprocally promoting each other in a representation preserves itself), or for producing its object. (*FI* 20:230-1)

Pleasure, according to this passage, is a mental state in which a representation "agrees with itself". Although this phrase is obscure, the idea seems to be that pleasure functions as a ground - i.e., a reason or a cause - for preserving itself, or for producing its object. The former is a pure aesthetic pleasure, where the mind's faculties are in harmony ("reciprocally promoting each other"), while the latter is probably a “practical pleasure”, that necessarily causes a desire for the object (*MM* 6:212).⁵⁴ So, all pleasures are behavioral dispositions. But while some pleasures are dispositions to preserve one's pleasant mental state (say, by staying put while observing a Monet painting so that the pleasant experience persists), others are dispositions to bring about an object whose representation is pleasant (as when my pleasure at the thought of a cup of coffee disposes me to get one). And we may reasonably infer that there is also a

⁵⁴ It is also possible, however, that in speaking of a pleasure that is a ground for producing its object, Kant is referring to a mental state of the faculty of desire (rather than to a mental state of the faculty of feeling that causes desire). This reading seems to be supported by Kant's saying that "a representation which produces the effort <conatum> for maintaining its state of representation <statum repaesentativum> is called *pleasure*, one which becomes the cause for the production of an object is called *desire*." (*LM* 28:675).

displeasure that causes an aversion of an object, such that we are disposed not to produce the object (as when my displeasure at the thought of conflict disposes me to avoid a colleague, or to resolve the conflict); and there is also a displeasure that disposes us to change our mental state (as when displeasure disposes me to look away from an ugly painting).⁵⁵ So, pleasure and displeasure, according to the above passage, are dispositions to preserve or change one's states, or to act in order to produce or not to produce their objects.

A similar dispositional characterization of pleasure and displeasure comes from the following passage:

The consciousness of the causality of a representation with respect to the state of the subject, **for maintaining** it in that state, can here designate in general what is called pleasure; in contrast to which displeasure is that representation that contains the ground for determining the state of the representations to their own opposite (hindering or getting rid of them). (*CPJ* 5:220)

Kant is saying here that pleasure (+) is the consciousness of the causality of a representation for maintaining the subject in its state, and displeasure is a representation that contains the ground for hindering a representation or getting rid of it. These characterizations of pleasure and displeasure do not exactly mirror each other, but the idea seems to be that pleasure (+) is a mental state in which we are conscious of a cause or a reason for preserving this state, while displeasure is a mental state in which we are conscious of a cause or a reason for changing our state. So, here too, pleasure and displeasure are understood as behavioral dispositions. And although Kant neglects to mention that pleasure (+) may dispose us to produce an object, this need not be taken to suggest that this characterization applies only to aesthetic pleasures, and not

⁵⁵ On my reading, *FI* 20:230-1 is compatible with there being a pure aesthetic displeasure, that disposes us to change our mental state. Ginsborg (2003, 175-77) and Guyer (2005, Chapter 6), in contrast, deny that there is such a thing as a pure displeasure in the ugly in Kant. That Kant thinks there is a pure aesthetic displeasure is suggested when he says "What pleases from agreement with the universal sense, that is beautiful, and if it displeases from the same ground, it is ugly." (*LM* 28:249).

to practical ones; for a disposition to preserve one's state may be understood in a broad sense, that includes producing the object that would preserve one's state.⁵⁶

The fact that Kant characterizes pleasure in phenomenological, evaluative, and dispositional terms, might lead one to think that he simply changed his mind on this matter. But this conjecture is implausible, given that Kant sometimes shifts between these characterizations within a short bit of text or a single lecture. Such a shift occurs in *FI* within a single paragraph: Kant characterizes pleasure as a disposition, saying that it is (or functions “as”) a ground for preserving one's mental state or for producing its object, and then hints at its felt subjective quality, saying that pleasure and displeasure are “felt, not understood” (20:230-2).

Kant is also recorded speaking about pleasure and displeasure as feelings, dispositions and evaluations within a single lecture. In a metaphysics lecture dated 1782-3, he says that “Pleasure is the representation of the agreement of an object with the productive power of the soul, and displeasure the opposite” (*LM* 29:894), employing an *evaluative* account. He then goes on to say that “pleasure and displeasure is a feeling of agreement and conflict or, what is the same thing, of the promotion or obstruction of life” (*ibid.*), which suggests that pleasure and displeasure are *feelings* or phenomenological qualities, or that they are representations of the promotion or obstruction of life.⁵⁷ And earlier in the lecture he says that “The feeling of pleasure

⁵⁶ Many other passages characterize pleasure and displeasure, or their sensuous species, in dispositional terms, without distinguishing between dispositions directed at one's mental state, and those directed at the object: “Pleasure is the inner determining ground of activity, the representation which determines the subject to activity.” (*LR* 28:1275, translation mine); “Every discomfort or pain requires us to leave our present condition, and this is its definition.” (*LA* 25:1316); “What directly (through sense) urges me to leave my state (to go out of it) is *disagreeable* to me - it causes me pain; just as what drives me to maintain my state (to remain in it) is *agreeable* to me, I enjoy it.” (*Anth.* 7:231); “Pleasure is when a representation contains a ground for being determined, for producing again the same representation, or for continuing it when it is there.” (*LM* 28:586). I take these characterizations to apply to both pure aesthetic and practical pleasures.

⁵⁷ Kant associates the feeling of life both with a free and regular play of our faculties, and with certain physiological changes: “When the blood and animal spirits are set into powerful motion, and when the sensation is more intense on one spot, then this feeling is an enjoyment, although life is itself thereby

is the ability of my power of representation to become determined by a given representation to its maintenance or promotion or avoidance" (*LM* 29:890), alluding to a *dispositional* conception.

The fact that he shifts between the three characterizations within a single lecture makes clear that these characterizations do not reflect conceptions held by Kant at different junctures in his career. Unless we are ready to say that Kant was confused about the nature of pleasure, we had better find a way to combine the evaluative, phenomenological, and dispositional components of pleasure into a unified account. How might such an account work?

One possible answer is that pleasure is an aggregate of three separate mental states, causally related to one another: an evaluative mental state causes its subject to experience a certain felt quality, which in turn causes a disposition to action. But if this is how the three components are related, it is unclear why their conjunction should constitute a single mental state called "pleasure" that originates in the faculty of feeling. It rather seems more likely that pleasure consists in a felt quality, produced by the faculty of feeling, while the evaluative and dispositional components, although necessarily connected with it, are not constitutive of pleasure itself - they rather belong to the faculty of cognition and desire, respectively.

But while this reading seems to have the advantage of maintaining a strict division of labor between the three mental faculties, it is problematic for two reasons. First, it conflicts with passages where Kant speaks of pleasure as having representational content (e.g., *CPrR* 5:9, *LM* 28:894) and as a ground for activity (*CPJ* 5:220). In those passages, pleasure is not described as *connected* to representations that have a certain content and to grounds for activity, but rather as itself a kind of evaluation and disposition. Second, the kinds of evaluation and disposition

impeded." (*LA* 25:559). He is explicit, however, that feeling is not a reliable means for tracking whether our life is actually promoted: "There can be an enjoyment which diminishes life, but increases the feeling.... Pain is not related to the proportion of ill, but the proportion of the feeling of ill." (*ibid.*).

involved in pleasure are such that they cannot originate in the faculties of cognition and desire. For the evaluation of pleasure is of an object's relation to the subject, and as such, this evaluation cannot be produced by the cognitive faculty; and at least some of the dispositions that pleasures consist in are different from the dispositions of the faculty of desire, as they are directed at preserving or changing one's state (rather than at producing or not producing an object).

The account of pleasure as an aggregate of three causally connected mental states is therefore implausible, insofar as it threatens to disintegrate it into components which belong in the other faculties. But if pleasure is not a mere aggregate of three separate mental states, it must somehow unify the evaluative, dispositional and feeling components into a single mental state. How can these three components hang together? The answer, I think, lies in the kind of evaluation that pleasure consists in. For, as we have seen, if pleasure represents the way the subject is affected or modified, it cannot be a "cold" objective representation or mental state - it rather must have a felt quality to it. And if pleasure (+) is a representation of an object's agreement with one's ability or intention to act, then it must dispose its subject to act - either in order to preserve its mental state, or in order to bring its object about. Conceived in this way, pleasure is not a mere conjunction of three distinct mental states; it is rather a single mental state that evaluates the object's relation to us, disposes us to action, and has a certain felt quality.

This account of pleasure is attractive, I propose, because it incorporates the three main functions theoreticians have traditionally assigned to pleasure. According to one traditional theory, pleasure is a simple feature of experience, that makes experiences good to the extent that it is present (Katz attributes to Locke something similar to this view, in Katz 2016, §1.1); others take pleasure to be evaluative (e.g., Helm 2002); and still others think it consists in a behavioral

disposition (Ryle 1949, Chapter IV).⁵⁸ Although these approaches are often formulated to the exclusion of one another, we have seen that all three functions appear in one form or another in Kant. And if what I have argued is correct, Kant's account successfully combines them into a unified, coherent account, such that they all belong to one mental state.⁵⁹

2.5. Two Alternative Accounts

I have argued that pleasure for Kant involves an evaluation, a disposition, and a felt quality. This account is similar to the “phenomenological model” Guyer thinks is sometimes at play in Kant because it takes pleasure to have a felt quality; and it is similar to Cohen's account in that it takes pleasure to represent something. In contrast with Guyer's phenomenological model, however, I do not think that all pleasures have the same phenomenological quality; and in contrast with Cohen, I do not think that pleasure represents by means of reflective judgment. The aim of this section is to explain why I disagree with them on these points.

Starting with Guyer, as mentioned earlier, he thinks that at least in some passages, Kant is committed to a *strong phenomenological model*, on which pleasure and displeasure are identical within their respective tokens, such that all pleasures feel the same, and all displeasures feel the

⁵⁸According to Katz, recent empirical findings about the intricate connections of pleasure with various physiological functions (e.g., Thayer 1989; Rosenkranz *et al*, 2003; Craig 2002, 2009, 2014) suggest that pleasure may be “more a syndrome of typically causally connected features than a simple or unified psychobiological phenomenon, such as would better fit philosophers’ penchant for simple kinds and simple explanations.” (Katz 2016, §4). Kant's account, as I have interpreted it, agrees with Katz's proposal insofar as it takes pleasure to involve several phenomena. But for Kant, these phenomena constitute one mental state, rather than a “syndrome” of causally connected features.

⁵⁹ As mentioned in the previous chapter, evaluation, motivation, and feeling are also the three functions philosophers have traditionally associated with emotions. Since on our assumption, some of the mental states of the faculty of feeling are emotions, at least some emotions in Kant combine the three functions traditionally associated with emotion.

same. Guyer's main new piece of textual evidence in support for this strong phenomenological model is the following:⁶⁰

The agreeable, the beautiful, and the good therefore designate three different relations of representations to the feeling of pleasure and displeasure, in relation to which we distinguish objects or kinds of representation from each other. The expressions appropriate to each of these, by means of which one designates the pleasure in each of them, are also not the same. **Agreeable** is that which everyone calls what **gratifies** him; **beautiful**, what merely **pleases** him; **good**, what is **esteemed, approved**, i.e., that on which he sets an objective value. (*CPJ* 5:209-10)

According to Guyer, this passage suggests that "the difference among three main cases of pleasure lies not in any quality of the feeling of pleasure, *which is the same in each case*, but in the way in which that feeling is related to, specifically caused by, its object, and in the status of the judgment about the object that may be made given the way in which it causes pleasure" (Guyer 2018, 147-8, italics mine). It is not immediately clear why Guyer thinks this passage suggests that the feeling of pleasure is "the same in each case" in the three kinds of pleasure it mentions (i.e., in the agreeable, beautiful and the good). But we might have a reason for thinking so if we were to read the phrase "the feeling of pleasure and displeasure" as referring to mental states of two different types, i.e., mental states of pleasure and displeasure. Read in this way, Kant is saying that the agreeable, the beautiful, and the good, designate three different relations of representations to one and the same *type of mental state*. Indeed, the definite article in the phrase "*the* feeling of pleasure and displeasure" suggests that the three cases of pleasure are in fact of the same kind or type, and so it makes sense that they would all feel the same.⁶¹

⁶⁰ Allison (1998) argues that Guyer is not warranted to attribute to Kant the strong phenomenological model on the basis of the evidence presented in Guyer (1997). I take my response to Guyer to advance on Allison's in that it examines Guyer's new evidence as well as his attempt to respond.

⁶¹ In his 1997, Guyer says that "Kant always refers to the feeling of pleasure in the singular, implying that, however various instance of pleasure may differ in origin or intersubjective validity, the nature of the

As we have seen earlier, however, the phrase “the feeling of pleasure and displeasure” is ambiguous. It could refer to (a) the mental faculty or susceptibility that enables us to experience pleasure and displeasure; or to (b) mental states of pleasure and displeasure (be they types or tokens). In this passage, however, “the feeling of pleasure and displeasure” has sense (a). For if Kant meant to speak here about the relation of “agreeable”, “beautiful” and “good” representations or objects to the kind of mental states they cause, he should have said these terms designate relations to the feeling of pleasure - not to the feeling of pleasure *and displeasure*. For the objects or representations to which we attribute these predicates (with the exception of the morally good)⁶² give rise to pleasure alone - and not to displeasure. It is therefore more reasonable to read Kant as saying here that the agreeable, the beautiful and the good all designate relations to one and the same *mental faculty*. Of course, all “positive” token mental states produced by the feeling of pleasure and displeasure are tokens of pleasure or of feeling, and in this sense, they are of the same type. But this does not entail that they all feel the same. To the contrary: Kant uses here distinct terms not only for the attributes “agreeable”, “beautiful” and “good”, but also for the kinds of pleasure that objects or representations with these attributes elicit in us (gratification / enjoyment, pleasantness, esteem / approval). This suggests that all pleasures *do not* feel the same. Rather, pleasure comes in three phenomenological “flavors”: gratification, pleasantness and esteem.⁶³

feeling itself is always the same.” (Guyer 1997, 103). Here, the reasoning is that because there is only one feeling of pleasure, all pleasures feel the same. But if “the feeling of pleasure” designates a mental faculty rather than a mental state of some kind, the fact that there is only one such faculty does not imply that all its instances feel the same.

⁶² This claim requires qualification because the thought of a morally good action gives rise to the feeling of respect for the moral law, which is a combination of pleasure and displeasure (*CPrR* 5:73, 78-9). But all other objects that have the attributes mentioned in 5:209-10, including the good which is merely useful (and not moral), provoke pleasure, not displeasure.

⁶³ The suggestion that the three types of pleasure differ in their felt quality may be thought implausible, insofar as it conflicts with Kant's claim that judgments of taste are fallible (*CPJ* 5:237). For if an error in a

Another passage to which Guyer refers as evidence for the strong phenomenological model is this:

However dissimilar representations of objects may be - they may be representations of the understanding or even of reason, in contrast to representations of sense – the feeling of pleasure by which alone they properly constitute the determining ground of the will (the agreeableness (*Annehmlichkeit*), the gratification (*Vergnügen*) expected from the object, which impels activity to produce it) is nevertheless of one and the same kind not⁶⁴ only insofar as it can always be cognized only empirically but also insofar as it affects one and the same vital force that is manifested in the faculty of desire, and in this respect can differ only in degree from any other determining ground. (*CPrR* 5:23)

Kant says here that the feeling of pleasure by which alone representations of objects constitute a determining ground of the will is of one and the same kind, and can differ only in degree from any other determining ground. This might be taken to suggest that feelings of pleasure cannot differ in their felt quality. But, as Allison notes in his response to Guyer (1997), in this passage, "Kant is not talking about pleasure *per se*, but rather about pleasure as a motivating factor (determining ground of the will). In other words, the claim is merely that any qualitative difference between pleasures is irrelevant in this regard, since what matters is merely the effect of the pleasure on the faculty of desire. And this is certainly far from a blanket denial of any such difference." (Allison 1998, 475).⁶⁵

judgment of taste can occur only by confusing pleasure in the agreeable or the good for pleasure in the beautiful, and if these three kinds of pleasure are phenomenologically distinct, errors in judgments of taste are impossible. One way to respond to this objection would be to offer an alternative explanation of errors in judgments of taste, for example, that such errors occur when we misrepresent an object as promoting our life (or as agreeing with our or ability or intention to act). Kant allows for such errors, when he says "There can be an enjoyment which diminishes life, but increases the feeling... Pain is not related to the proportion of ill, but the proportion of the feeling of ill." (*LA* 25:559).

⁶⁴ This word does not appear in Guyer's citation, but I do not think this affects his reading.

⁶⁵ Similarly, Rueger argues that it is only with respect to the faculty of desire that all pleasures "reduce to degrees of agreeableness." (Rueger 2020, 131).

In what appears to be an attempt to respond, Guyer argues that Kant makes clear elsewhere (*MM* 6:212) that there is no qualitative difference between pleasure in moral and non-moral cases of motivation (Guyer 2018, 153). But although in the passage Guyer is referring to Kant does not mention a qualitative difference between moral and non-moral pleasure, neither does he say that there is no such difference.⁶⁶ Kant rather distinguishes there between desire in the narrow sense, which occurs when a determination of the faculty of desire "is caused and therefore necessarily *preceded* by... pleasure", and intellectual pleasure, which "can only follow upon an antecedent determination of the faculty of desire" (*MM* 6:212). This does not preclude there being a qualitative difference between a feeling of pleasure that precedes desire (which Guyer calls "non-moral pleasure"), and a feeling of pleasure that follows upon it ("moral pleasure"). But more importantly, in *MM* 6:212 Kant is considering pleasures as *motivational forces*. And so, even if moral and non-moral pleasures are qualitatively identical in this respect, they may be qualitatively different in other respects (and so, this passage does not address Allison's point). Granted, the way a pleasure feels cannot ground cognition, and this is why Kant says pleasure and displeasure cannot be explained more clearly in themselves (*MM* 6:212); but there may nonetheless be a felt quality which varies between different types of pleasures.

But although I think that Guyer's evidence does not support the view that all pleasures *feel* the same, I have argued that they *represent* the same thing, namely, agreement and disagreement with one's ability or intention to action. Cohen, in contrast, contends that different

⁶⁶ Elsewhere, however, Kant says that in judging the influence of the agreeable it is only a matter of the mass or intensity of the agreeable sensation; in contrast, the absolutely good is distinguished by the modality of a necessity resting on concepts *a priori* (*CPJ* 5:267). This suggests that although there is no qualitative difference between tokens of pleasures in the agreeable, i.e., enjoyments, there is such a difference between the pleasure of enjoyment and the moral feeling of pleasure in the absolutely good (that is, these two types of pleasure qualitatively differ).

feelings have different intentional objects. But on Cohen's view, feelings in Kant "have a form of 'derived-intentionality' that stems from their relationship with judgment." (Cohen 2020, 430). Cohen presents this original contribution to the ongoing debate on the intentionality of Kantian feelings as a middle-ground position between Guyer's (Guyer 1997) non-intentionalist and Allison's (Allison 1998) intentionalist reading of Kantian feelings, arguing that although they are not intentional in and of themselves (they do not have intrinsic / original intentionality, Cohen 2020, 435), Kantian feelings nonetheless have a "derived" or "derivative" intentionality. In a footnote, she refers to Byrne's definition of derivative intentionality: "A thing has derivative intentionality just in case the fact that it represents such-and-such can be explained in terms of the intentionality of something else; otherwise it has original intentionality." (Byrne 2006, 408). Accordingly, I take Cohen's claim that Kantian feelings have derived intentionality to mean that the fact that such feelings represent can be explained in terms of the intentionality of reflective judgments about them.

But as we have seen, there are passages where Kant speaks of pleasure as a representation or consciousness of a certain relation, which suggest that pleasures are intrinsically intentional. Kant says, for instance, that "Pleasure is the *representation of the agreement of an object or of an action with the subjective conditions of life*" (CPrR 5:9 fn.); that pleasure is "the representation of the agreement of an object with the productive power of the soul" (LM 29:894); and that it is "the consciousness of the agreement of an object with the productive power of imagination of our soul." (LM 29:891). These passages speak of pleasure as a representation or consciousness of something, and there is nothing in them to suggest that pleasure represents by means of reflective judgment.

It might be thought that feelings cannot have intrinsic intentionality because they have no objective reference.⁶⁷ But from the fact that feelings are subjective, and do not represent *objects* (or, more accurately, objects as they are independently of the way they affect us), it does not follow that they have no representational content at all, and must await reflective judgments to gain their intentionality. If what I have argued is correct, feelings are about objects' agreement and disagreement with the subject's ability or intention to act.⁶⁸ And so, although they do not tell us anything about the object as it is independently of its impact on us, they are nevertheless about something, independently of any judgments we might make about them.

2.6. Lingering Objections

If what I have argued is correct, Kant's account of pleasure incorporates an evaluative, phenomenological and dispositional component. Given that none of the alternatives discussed in 2.1 combines all three elements, their proponents would probably object that my reading combines some elements that do not belong in Kant's account of pleasure. I cannot discuss all these worries here, but I will respond to those I take to be the most serious.

First, it might be objected that a disposition or motivation cannot be essential to pleasure because pure aesthetic pleasures do not cause desires for their objects, and so they do not

⁶⁷ Cohen seems to be motivated by this worry when she says she is sympathetic with Guyer's view that "it is hard to see how feelings can have intentionality in and of themselves"; and she adds that "insofar as in feeling 'nothing at all in the object is designated' (CJ 89 [5:203]), it is not clear how feelings can refer to objects." (Cohen 2020, 435). It is not clear, however, why Cohen thinks that having no objective reference prevents feelings from having intrinsic intentionality.

⁶⁸ In his 2018, Guyer writes: "I sympathize with the queasiness about the idea of an entirely opaque sensation of pleasure, but have never been able to understand what sort of mental state a genuine sensation that yet has determinate intentional content is supposed to be – as a card-carrying Kantian, my view has been that determinate intentional content requires concepts and not just sensation." (Guyer 2018, 164). Guyer thus takes the non-conceptual nature of pleasures as preventing them from having determinate intentional content. Assuming that he is correct in that pleasures do not involve concepts, this would prevent them from being *cognitions* or judgments in Kant's usual sense (A68/B 93, B141). It is not clear, however, why this should deprive them of determinate intentional content.

motivate. And if this is the case, then my account is flawed insofar as it has the implausible consequence that all pleasures - even pure aesthetic ones - motivate. Or, as Cohen puts it, the existence of the feeling of the beautiful, which is not necessarily connected to desire, shows that "the essential function of feeling cannot be to motivate." (Cohen 2020, 432).

In response, I think it is important to note that while pure aesthetic pleasures do not necessarily cause a *desire* for their objects, this need not imply that they do not motivate. For in his "transcendental definition" of pleasure (+), Kant says that it is "a ground, either merely for preserving this state itself... or for producing its object." (*FI* 20:230-1). But this definition is meant to apply to all pleasures (+), whether they are accompanied by sensation, reflection, or a determination of the will, and so it should apply to aesthetic pleasures (accompanied by reflection). Aesthetic pleasures (+) are of course not grounds for producing their objects, but they are nonetheless grounds for preserving one's state, such that when we experience them, we have a reason to preserve our state and are motivated to do so. For example, in experiencing pure aesthetic pleasure while observing a piece of Monet's *Water Lilies*, I am not motivated to produce its object - neither the painting itself, nor the flower garden depicted in it. But I am nonetheless motivated to preserve my pleasant state, for example, to concentrate on the artwork, and to shut out anything that might break my concentration. And so, although this pleasure does not cause a desire for its object, it does motivate to preserve itself.⁶⁹ And if pure aesthetic pleasures do motivate, then my account is not flawed for taking motivation to be essential to all pleasures.⁷⁰

⁶⁹ Since not being motivated that my state persists seems tantamount to being *indifferent* to my mental state's persistence, the idea of a pleasure that involves no motivation to preserve itself borders on the non-sensical. In what sense could my experience be pleasant, if I am indifferent as to whether it persists?

⁷⁰ Is the motivation to preserve one's state in pure aesthetic pleasure not the same as a desire? I think the answer to this question is "no" because such a motivation is a feeling, and as I will argue in the next

This first objection was that motivation or disposition to action is not essential to pleasure, and should be excluded from its account. But one might also object that there is evidence that a behavioral disposition is *the only thing* that is essential to pleasure, and so the evaluative and phenomenological components should be left out of my account. Such evidence, the objection goes, is found in passages where Kant characterizes or even defines pleasure and displeasure or pain in dispositional terms, and does not mention that they involve felt qualities or evaluations. For example, as we've seen, Kant says that "The consciousness of the causality of a representation with respect to the state of the subject, for **maintaining** it in that state, can here designate in general what is called pleasure" (*CPJ* 5:220). The fact that Kant does not mention here that there is a felt quality or an evaluation involved in pleasure suggests either that these features are not involved in pleasure at all, or, if they are involved, they are not essential to it. Moreover, in one lecture, Kant is recorded saying that pain is defined as that which "requires us to leave our present condition" (*LA* 25:1316). But if pain is *defined* in dispositional terms, then either it involves no felt quality and no evaluation, or the felt quality and evaluative components are inessential to it and should be excluded from its definition.⁷¹

My response to this objection is that although these passages suggest that pleasure may be characterized and even defined solely in dispositional terms, Kant's view nonetheless appears to be that we cannot provide a *proper* explanation or definition of pleasure; and if this is the case, it is possible that his explanations of pleasure do not include all its essential features. That Kant thinks we cannot provide a proper explanation of pleasure is strongly suggested by passages

chapter, feelings differ from desires in that the former are directed at the subject's state while the latter are future-directed.

⁷¹ Guyer appeals to these two passages to support his claim that "we can suppose that at least sometimes he (Kant - U.E.) thought that the dispositional model of pleasure and pain was adequate." (Guyer 2018, 158).

where he warns that we cannot explain pleasure by itself (*MM* 6:211, *FI* 20:230-2); and a similar, albeit weaker warning, appears just before Kant “defines” pain as that which requires us to leave our present condition (*LA* 25:1316): “it is all the more difficult to provide distinct concepts of it (of satisfaction, or of the feeling of pleasure and displeasure - U.E.).” (*LA* 25:1316). One way to make sense of these warnings is by appeal to the idea that a proper definition, in one sense this term has in Kant, determines the content of a concept by means of its marks,⁷² which means that phenomena with no marks cannot be properly defined or explained. And since a mark is the part of the representation that allows us to cognize the represented object (16:298), pleasures - as subjective, non-cognitive mental states - have no marks, and cannot be properly defined. But if definitions of pleasure are not proper definitions (or not explanations of pleasure “by itself”), then the dispositional definitions that Kant does provide need not be taken to include everything that is essential to pleasure; and so, the phenomenological and evaluative components mentioned in other passages should not be excluded from our account of pleasure.⁷³

Finally, one might argue that there is philosophical (rather than textual) reason not to attribute to Kant an account on which pleasure and displeasure have a felt quality. Guyer offers such a reason when he says:

To be sure, some cases of pleasure must involve distinctive sensations, for there is a characteristic way or range of ways that a good Bordeaux tastes, and a different way that a good Burgundy tastes, and each is enjoyable; but it is less plausible that there is a distinctive feeling of pleasure, whether always the same or not, *in addition to* the characteristic Bordeaux taste and Burgundy taste. So perhaps there is room for an account of pain that combines a phenomenological and a dispositional aspect, but it is less clear that this will be so for an account of pleasure. (Guyer 2018, 163)

⁷² Messina and Sturm 2015, 559; A712-738/B740-766.

⁷³ It should be noted, however, that if my understanding of Kant's restriction on explaining pleasure is correct, then my account of pleasure is not a proper definition of it: it does not determine the concept of pleasure by means of its marks.

One way to read Guyer's claim here is as appealing to introspection, such that when we reflect upon our pleasure in a Burgundy, we cannot find any feeling aside from the wine's characteristic taste. Or, to pick up on Guyer's other example, when reflecting upon his experience while observing Vermeer's *View of Delft*, he cannot recall any distinct sensation of pleasure - all he remembers is the disposition to stay put, wanting the experience not to end. Another option is that Guyer is appealing here to considerations of theoretical parsimony: admitting that in addition to objective sensations such as the wine's taste or the picture's color, there is another set of mental states called "feelings" would be multiplying mental entities without necessity.

If Guyer is appealing to introspection, then he is right insofar as it is often *difficult* to detect or recall a distinct feeling of pleasure in addition to its associated representation. But given that introspection is fallible, appeal to it provides no decisive objection to the claim that pleasure has a feel to it (be it uniform across its types or not). If, however, Guyer is appealing to considerations of theoretical parsimony, then as we have seen, although Kant sometimes speaks of pleasure as a representation (e.g., *CPrR* 5:9 fn.; *LM* 29:894), at other times he speaks of it as an aspect of a representation (*CPJ* 5:189, *MM* 6:212 fn.).⁷⁴ So, considerations of introspection do not allow us to exclude the possibility that pleasure has a felt quality; and if it does have such a quality, this need not involve introducing a set of mental states distinct from those that give rise to them.

I therefore submit that these textual and philosophical considerations should not prevent us from attributing to Kant the suggested tripartite account of the mental states of the faculty of

⁷⁴ However, as we have also seen, if pleasure is merely an aspect of another representation, it probably cannot have an independent representational content. In other words, parsimony in the realm of mental states may come at the cost of the intentionality of pleasure.

feeling. On this account, such mental states are evaluations of objects' agreement or disagreement with our ability or intention to act, that have a felt quality, and dispose us to preserve our state or change it (or to produce or not to produce their objects). Assuming, as we have done, that some of the mental states we call "emotions" are associated by Kant with the faculty of feeling, the result is that some of the mental states we call "emotions" have this tripartite structure, according to Kant. We can now turn to the other set of mental states pertinent for reconstructing Kant's theory of emotion, namely, those he associates with the faculty of desire.

Chapter 3. Kantian Desires: A Holistic Account

3.1. Between Pleasure and Action

With this tripartite account of the mental states of the faculty of feeling under our belt, we can now proceed to examine the nature of the other set of mental states germane to understanding emotions in Kant, namely, those he associates with the faculty of desire, which I will henceforth refer to as “desires”.⁷⁵ By now it should already be clear that desires in Kant are closely connected to *pleasure* because as we have seen, desire and aversion always involve pleasure and displeasure (*MM* 6:211); and Kant also refers to desire itself as a kind of pleasure or satisfaction (for example, when he says that "Desire is the satisfaction [*Wohlgefallen*] in the existence of the object." *LA* 25:1334). But desire is also closely connected to *action*, as indicated by Kant's characterization of the faculty of desire as "the faculty to be, by means of one's representations, the cause of the objects of these representations" (*MM* 6:211), and by his characterization of desire itself as "the active power of the self-determination of the actions of thinking beings." (*LA* 25:577).⁷⁶

⁷⁵ Although I will sometimes refer here to desires, inclinations, passions and instincts as “mental states” and “determinations”, these last two terms are to be understood in a broad sense, that can refer not only to mental events keyed to particular moments, but also to dormant dispositions or tendencies to experience certain mental states.

⁷⁶ The connection between the faculty of desire and action is also suggested by Kant's characterization of life as "the faculty of a being to act in accordance with laws of the faculty of desire." (*CPvR* 5:9 fn.). Although it is not immediately clear how the faculty of desire and the faculty of life are related, one option is that the faculty of desire legislates for the faculty of life which executes its orders, in a way that is analogous to the way some commentators conceive of the relation between the faculty of desire *qua* will and *qua* power of choice (e.g., Allison 2020, 451). If this is correct, then God, to which Kant ascribes a faculty of desire (*LR* 29:1059), is “alive” insofar as he has the faculty to act according to the laws of the faculty of desire (he is a spiritual living being, *LM* 28:248). Although God probably cannot *fail* to act according to these laws, human beings can (for example, cases of weakness of will may be thought of as failures of the faculty of life to follow the orders of the faculty of desire), which accounts for the need to distinguish between the faculty of life and of desire. Moreover, it is important to note that since non-rational animals have a faculty of life, they too have the faculty to act in accordance with laws of the faculty of desire, but these laws are natural and necessitating for them. Another way of understanding the relation between these two faculties is that the faculty of life is a faculty of desire *of a certain kind*, namely, one that requires physical movement to bring about represented objects. In contrast, the faculty of

Commentators tend to focus on one of these components as essential to desire, and so their accounts may be classified into “feeling-based” and “action-based” accounts. Wood offers an account of the former kind, saying that “To *desire* an object (or state of affairs) is to have a *representation* of it accompanied by a feeling of *pleasure*. (*Aversion* is a representation accompanied by a feeling of displeasure.)” (Wood 1999, 50). A slightly different version is defended by Grenberg, who argues that “insofar as the representation of a feeling⁷⁷ includes representation of the pleasure (or potential pleasure) to be taken in the existence of a particular object or state of affairs, one can be said to have a ‘desire’ for the object in question; insofar as it includes pain,⁷⁸ one has an aversion to it.” (Grenberg 2001, 162). On Wood’s account, desire (+)⁷⁹ is a representation of an object or state of affairs accompanied by a feeling of pleasure, and aversion is such a representation accompanied by a feeling of displeasure. On Grenberg’s account, desire (+) occurs whenever we feel or expect to feel pleasure in the existence of an object or state of affairs, while aversion occurs whenever we feel or expect to feel displeasure in the existence of an object or state of affairs.⁸⁰

desire in some beings does not require physical movement. For instance, God’s faculty of desire can produce objects by merely representing them. If this is the case, God is not a living being (as he has no faculty of life).

⁷⁷ Although she uses the phrase “the representation of a feeling”, Grenberg’s view is that feelings are “connected to and evoked by representations” (Grenberg 2001, 162), which gives them what she calls “nascent intentionality”; and she says that “it is not clear whether Kant considers feeling itself to come under the genus of ‘representation’.” (Grenberg 2001, 164, fn. 28).

⁷⁸ Grenberg is using “pain” as the opposite of pleasure (+), synonymously with “displeasure”. As mentioned earlier, however, pain has a technical sense in Kant: it is a sensuous displeasure through sense (*Anth.* 7:230).

⁷⁹ Recall that a “+” sign indicates that the mental state it follows is a “positive” instance.

⁸⁰ I classify these accounts as “feeling-based” because they take feeling to be essential to desire. However, neither Grenberg nor Wood ignore the connection between desires and action. In Grenberg, this connection is evident in her claim that the nature, course and relative strength of desires constitute an agent’s drives (Grenberg’s rendition of Kant’s *Triebfeder*) to action, and are, all else being equal, “what determines action.” (Grenberg 2001, 163). In Wood, this connection comes up in his discussion of the “incorporation thesis”, the point of which he takes to be that “impulses or desires determine action chiefly by serving as incentives for the adoption of maxims.” (Wood 1999, 53). In correspondence, he has

In contrast, Frierson defends an “action-based” account, on which what is essential to desire is not a feeling of pleasure of some sort, but rather a causal connection to action. On his account, "a desire is *defined* as a representation that leads to action... And when desire is taken in this sense, there are no actions that are not preceded by and caused by desires, and virtually no desires that do not lead to actions (in the absence of external impediments)." (Frierson 2014b, 55). On this “action-based” account, desire is a representation that causes action, provided there are no external impediments.

The aim of the present chapter is to provide an account of desire that combines what these accounts get right, and solves the problems that inflict them. The “action-based” account, I will argue, holds for desires as volitions, but leads to denying that inclinations and wishes are proper desires (which conflicts with the text); the “feeling-based” account is correct insofar as desires involve pleasures, but does not explain how - if it all - desires differ from the practical feelings of pleasure they are causally connected to. On my proposed alternative, desires are motivations or dispositions to action, but need not bring it about, and inclinations and wishes are desires in this sense too; and desires are necessarily connected with practical pleasures but differ from them in that the former are normally directed at bringing about future objects by means of their representations (which allows them to motivate action immediately), while the latter refer to the subject's affective state (and so they must cause desires if they are to produce action).⁸¹

clarified that the above citation (i.e., Wood 1999, 50) was meant to provide the necessary conditions on a mental state's being a desire.

⁸¹ Wuerth argues that desires in Kant, qua impelling causes, "are representations of objects, coupled with satisfaction or dissatisfaction... urging us toward a choice of action" (Wuerth 2014, 227). Our accounts are similar in that we both take desire to have a pleasure (or satisfaction) component and a dispositional component that need not cause action. Our accounts differ, however, in that I distinguish between practical feelings and desires in terms of their temporal directedness, and explain why the evidence does not support the view on which desire causes action in the absence of external impediments.

I start by looking at Kant's characterization of desire in *Anth.*, and argue that it applies to sensuous and intellectual desires, be they habitual or not (3.2), and that desires, except for certain wishes, are directed at *future* objects (3.3). I proceed to probe into this characterization, and interpret it such that desire is a motive or disposition to action, which can but need not bring it about. In contrast to Frierson's, this reading does not require denying that inclinations and wishes are proper desires (3.4). I then look at passages where Kant speaks of desire as connected with pleasure and as a pleasure in an object's existence, and I propose to read them such that desires are causally connected with practical feelings, but differ from them in that the former are normally future-directed while the latter are not (even when connected with representations of future objects). It is the future-directedness of non-wishful desires that allows them to motivate actions immediately (3.5). I end by responding to some objections (3.6), and noting that this account anticipates contemporary holistic theories of desire that combine dispositional, evaluative, and hedonic components (3.7).

3.2. Desire and Its Species

In Kant's major ethical writings, we hear quite a lot about the faculty of desire and one species of its mental states, namely, inclination, but little is said about the genus to which inclinations and the other mental states associated with the faculty of desire belong. Fortunately, Book III of Kant's *Anth.* opens with the following helpful characterizations of desires, inclinations, wishes and passions, which may be used to launch our discussion:

Desire (Begierde) (appetitus) is the self-determination of a subject's power through the representation of something in the future as its effect [*als einer Wirkung derselben*].⁸² Habitual sensible desire is called *inclination*. Desiring without

⁸² The Cambridge translation reads "as an effect of this representation". The translation used by Kain, in contrast, reads "as an effect of that power" (Kain 2003, 234). But *einer Wirkung derselben* is ambiguous, and may refer to (a) the effect of the self-determination of the subject's power; (b) the effect of the

exercising power to produce the object is *wish*. Wish can be directed toward objects that the subject himself feels incapable of producing, and then it is an *empty* (idle) wish... Inclination that can be conquered only with difficulty or not at all by the subject's reason is *passion*. (*Anth.* 7:251)

As I read this passage, the first sentence provides a characterization of the genus “desire”, and its remainder lays out a taxonomy of the different species of desire, which include inclination, wish, empty wish, and passion. That all these are species of desire is strongly suggested by the fact that Kant refers to all of them as desires: inclination is habitual sensible *desire*,⁸³ and passion is desire by virtue of being a kind of inclination (one that cannot be easily conquered by reason). In the case of wish, Kant uses the substantive-infinitive form “desiring” (*Das Begehren*) rather than the count noun “desire” (*Begierde*) to refer to it. But desiring is naturally understood as the mental activity one is engaged in when one has a desire, so that wishing is an activity of desiring, and to have a wish is to have a certain kind of desire.⁸⁴ In other passages, Kant adds to this taxonomy of the mental states of the faculty of

subject's power, or (c) the effect of the representation of something in the future. My modified translation (“as its' effect”) preserves this ambiguity, but I will argue below that (c) is the best reading.

⁸³ On an alternative reading, Kant is contrasting inclination, which is a habitual desire, with desire, which is necessarily non-habitual; and if this is the case, inclination is not a species of desire, and the characterization of desire does not apply to it. But this suggestion seems to me misguided because if inclination were not a species of desire, there would have to be some genus to which both inclination and desire belong as species. The fact that Kant does not provide such a generic term, but rather opens his discussion with a characterization of desire, suggests that desire is the genus to which all the mental states of this faculty - habitual or not - belong. It might be objected that the fact that desire is characterized as self-determination of the subject's power suggests that it must occur at a particular moment and cannot be a habit, which is a disposition. But “self-determination” may be understood as a disposition or motivation to action, in which case it *can* be a habit. I will argue in 3.4 that this is how we should understand it.

⁸⁴ Some passages (*MM* 6:211, 23:262) suggest that *Begehren* is the act of the faculty of desire which necessarily brings action about (also called “volition”, *LA* 25:1344), while *Begierde* is a possible ground for such an act. This is not how Kant uses these terms in *Anth.* 7:251, however. For there, the phenomenon that Kant calls *Begierde* (i.e., the self-determination of the subject's power) comes closer to an act of the faculty of desire than the one he calls *Begehren* (i.e., wish, which is desiring without exercising power to produce the object). So, there is no reason to think that *Begehren* here means something stronger than *Begierde*, nor that it must involve an act of the faculty of desire that leads to action.

desire two other species: propensity, which is "The subjective *possibility* of the emergence of a certain desire, which *precedes* the representation of its object" (*Anth.* 7:265); and instinct, which is "the inner *necessitation* of the faculty of desire to take possession of this object before one is even acquainted with it is [*ehe man ihn noch kennt*]." (*Anth.* 7:265; cf. *Rel.* 6:29 fn.). A propensity is probably best understood as a *potential* rather than an actual desire,⁸⁵ and some argue that the same holds for instincts.⁸⁶ But independently of whether propensity and instinct qualify as actual desires, inclinations, passions and wishes are all species of desire.

“Desire”, however, is an ambiguous term in Kant. “Desire in the narrow sense” is “that determination of the faculty of desire which is caused and therefore necessarily *preceded* by such

⁸⁵ The characterization of propensity in *Anth.* 7:265 is actually ambiguous between: (1) a potential for *acquiring* a desire before one has had a representation of the desired object, such that propensity is eradicated once the desire is acquired; (2) a potential for *experiencing* a desire, that persists after acquiring the desire (since the potential is there whenever we do not actually desire the object). Kant's example of the propensity for intoxicants suggests he has the first sense in mind ("...all savages have a propensity for intoxicants; for although many of them have no acquaintance at all with intoxication, and hence absolutely no desire for the things that produce it, let them try these things but once, and there is aroused in them an almost inextinguishable desire for them.", *Rel.* 6:29 fn.). However, the fact that he calls a sense-free inclination *propensio intellectualis* (*MM* 6:213) suggests that inclinations can function as propensities in the second sense, i.e., as potentials for experiencing occurrent inclinations.

⁸⁶ Wood takes all instincts to be potential desires, as he says that instinct is "the predisposition to be pleased by something even before one possesses it" (Wood 2014, 135). Frierson, in contrast, allows that instincts in Kant can be either actual and potential desires, as he says that instincts are usually predispositions to desire, triggered by cognition of the desired object, but grants that in some cases, "one can feel an instinct without a direct cognition of the object of that instinct." (Frierson 2014b, 67, fn. 28). Prima facie, there is reason to try and avoid attributing to Kant the view that instincts are actual desires: if an instinct is an actual desire for an object we have not been acquainted with, it seems to presuppose an innate representation of the object, which is both implausible and incompatible with Kant's rejection of innate representations (2:392-3; 8:221). In his lectures, however, Kant makes clear that instincts *are* actual desires, saying of instinct and inclination that "Both are actual desires." (*LA* 25:584), and that "Instinct is an actual desire, but without clear cognition of the object." (*LA* 25:1339). I therefore think that Kant's view is that instincts are actual desires that precede *acquaintance* with their objects, but do not precede *representations* of their objects. This is possible because to be acquainted (*kennen*) with an object in Kant's technical sense, involves a cognitive capacity more complex than the one involved in perceiving or having a representation of an object, i.e., the capacity to distinguish the object from its surroundings (see *LL* 24:135, 9:65). So, the idea that there are desires that precede acquaintance with their objects does not commit Kant to innate representations.

(practical - U.E.) pleasure" (*MM* 6:212).⁸⁷ But pleasure can also be regarded as caused by desire, and Kant mentions that there is an intellectual pleasure, namely, a pleasure which "can only follow upon an antecedent determination of the faculty of desire" (*ibid.*). So, there are two kinds of desire, or two kinds of "determinations" of the faculty of desire, according to Kant: "desires in the narrow sense", caused by pleasures, and desires that cause pleasure (called "sense-free inclinations", when they are habitual).⁸⁸ And so, "desire" *simpliciter*, not in the narrow sense, refers to both sensuous and intellectual determinations of the faculty of desire. Kant uses "desire" in this broad sense when he says in a lecture, for example, that "All desires are, like pleasure and displeasure, intellectual or sensitive." (*LM* 29:894; cf. 28:587).⁸⁹

⁸⁷ While one way of understanding desire and the practical pleasure it is causally connected to is as two distinct mental representations, Höwing proposes that they are two aspects or properties of the same representation (Höwing 2013, 30). I have no objection to this proposal, as long as we keep in mind that the faculty of desire provides access to one of these aspects or properties, and the faculty of feeling provides access to another.

⁸⁸ It might be objected to Kant's taxonomy of desires into those that cause and those caused by pleasure that some desires seem to be disconnected from feelings of pleasure altogether. For instance, when, on a cold winter day, I suddenly crave a cup of hot soup, my desire need not cause me a feeling of pleasure, and it need not be caused by such a feeling. I see three ways of defending Kant from this objection. First, it could be argued that although I am not *conscious* of a feeling of pleasure that brought about my desire for soup, a non-conscious feeling of pleasure (in the non-conscious representation of soup) must have occurred if I now desire it. This, however, conflicts with the evidence that feelings are necessarily conscious. Second, if pleasure is understood as a ground for preserving one's state or producing an object (*FI* 20:230), then my desire, as a ground for getting soup, involves or consists in pleasure. Finally, if rational desires that cause pleasure need not be moral ones, we might say that my desire for soup is rational in the sense that it was caused by my thought that having soup would keep me warm, and this desire in turn caused pleasure. However, Kant's view seems to be that desires that cause pleasures are moral desires that depend on "pure rational principles alone" (*MM* 6:212-3). This is the view attributed to him by Baum (2006, 135) and Höwing (2013, 31-2).

⁸⁹ Wilson argues that desires in this narrow sense are "episodes of desiring activity. They are occurrent states, akin to urges, itches and yens."; he adds, however, that "Kant does not think of inclinations as occurrent desires or episodes. Instead, he thinks of them as dispositions to experience such episodes." (Wilson 2016, 214). Wilson's view is therefore that inclinations are *not* desires in the narrow sense because they are not occurrent states akin to urges. But the distinction between desire in the narrow sense and desire not in this sense is not between habitual and non-habitual desires, but rather between sensuous desires (caused by pleasure), and desires in a generic sense (some of which are caused by pleasure, others causes of pleasure); and inclinations (that are not intellectual) are sensuous desires. Granted, Kant's characterization of desire in the narrow sense in *MM* 6:211 as a determination caused by pleasure might be taken to suggest that desire in this sense must be an occurrent mental state. But as pointed in fn. 83, a desire could be a determination of the faculty of desire in the sense that it is a disposition to desire certain

Which of these senses is at play in the first sentence of *Anth.* 7:251? Prima facie, it might be thought that “desire” here means “desire in the narrow sense” because this passage is taken from *Anth.*, understood by some to be providing an empirical psychology governed and causally determined by natural laws.⁹⁰ And this reading might seem to gain further support from the fact that the species of desire mentioned here include only sensuous desires, i.e., inclination (in its sensuous guise) and passion. But even if *Anth.* does provide a deterministic empirical psychology (which is a contested claim, see fn. 90), this need not imply that “desire” is used here in the narrow sense; for such a deterministic account could apply to intellectual desires as well.⁹¹ And although intellectual desires are missing in the published *Anth.*, in his lectures on anthropology, Kant speaks of rational desires that have ends set by reason (*LA* 25:579), of the higher faculty of desire or the will (*LA* 25:438; or “upper” faculty of desire, “where reason is incentive” *LA* 25:1334), and of intellectual desires that arise from motives (*LA* 25:274). So, the exclusion of intellectual desires from *Anth.* is not because they are incompatible with Kant's theoretical framework in the field of anthropology; it is better explained, I propose, as due to his focusing on obstacles to moral action.

Moreover, there are several reasons to think that “desire” in the first sentence of *Anth.* 7:251 is used in its broad sense, that refers to both sensuous and intellectual ones. First, the characterization of desire there (“the self-determination of a subject's power”, etc.) seems to

objects; and such a disposition could be caused by pleasure in the sense that it is triggered by it. If this is the case, inclinations can be desires in the narrow sense, as the latter need not be occurrent states.

⁹⁰ For example, Frierson argues that “Empirical psychology and anthropology both investigate empirically accessible and hence deterministic laws of human behavior” (Frierson 2014b, 47). Wood, in contrast, holds that Kant's empirical anthropology “always proceeds on the fundamental presupposition that human beings are free, and throughout it interprets the empirical observations it makes on the basis of this presupposition.” (Wood 2003, 44).

⁹¹ This seems to be Frierson's view, as he provides an account of moral motivation in Kant within the framework of an empirical psychology governed by causally necessitating laws (Frierson 2014b, 122-7).

be neutral between sensuous desires caused by pleasure, and intellectual desires that cause it. Second, if “desire” there meant “sensuous desire”, it would be redundant to note in the next sentence that inclination is a sensuous (“sensible”, on the Cambridge translation) desire, for this would be implied by the fact that it is a desire. Finally, in one lecture, Kant uses the Latin term *appetitio per motiva* for intellectual desires (*LM* 29:895), and in another he speaks of *Appetitiones* which are intellectual (*LM* 28:587). Since in 7:251 Kant glosses “desire” with the Latin *appetitio*, he is probably speaking of desire in the same sense, such that it may be intellectual.⁹² I therefore submit that Kant's characterization of “desire” here applies to both intellectual and sensuous desires, in their habitual and non-habitual guises.

3.3. Desire and Foresight

Assuming that “desire” in the first sentence of *Anth.* 7:251 refers to both sensuous and intellectual desires, be they habitual or not, we can proceed to examine its characterization as “the self-determination of a subject’s power through the representation of something in the future as its effect”. And we can unpack this dense characterization by dissecting it into two components: (1) the self-determination of a subject’s power; (2) done through the representation of something in the future as its effect.

Starting with the latter, desire, according to it, occurs by means of representing something in the future, where “something” is to be understood as an object in a broad sense, that may include physical objects (as when Kant speaks of an inclination for Brandy, *Anth.*

⁹² Giordanetti shares my view that *Begierde* in this passage means both sensuous and intellectual desires, for he comments that in this wide sense of *Begierde*, “every exercise of the faculty of desire is a desire or a desiring (*eine Begierde oder ein ,Begehren'*)” (Giordanetti 2015, 233, translation mine). It is interesting to note that in Leibniz, *appetitio* also need not be sensuous; it rather refers to a tendency or striving toward a new perception (Kulstad and Carlin 2020, §4). Kant's use of the same Latin term here suggests that he is using it in a similar manner.

7:269; or the desire for intoxicants, *Rel.* 6:29 fn.), actions (as an inclination for drinking, gambling, or hunting, *Anth.* 7:269), and states of affairs (for passions aim at the ability to satisfy inclinations, by having honor, power or money, *Anth.* 7:270).⁹³ For example, when I desire ice cream, my power is “self-determined” by means of a representation of ice cream (or my having it) as “its' effect”. “Its” here is ambiguous, however, and could refer to: (1) the representation of something in the future (ice cream or having it, in our example); (2) the self-determination of the subject’s power; or (3) the subject's power. Assuming that the subject's power must be self-determined if it is to cause a desired object - such that (3) either involves this self-determination, in which case it is the same as (2), or it doesn't, in which case it can be eliminated - we can narrow it down to (1) and (2). And (1) seems more plausible because “the representation of something in the future” is closer to the pronoun “its”.

This reading gains further support from Kant's characterization of the faculty of desire as "the faculty to be, by means of one's representations, the cause of the objects of these representations." (*MM* 6:211). According to this passage, the faculty of desire enables us to cause the objects of our representations by means of our representations; and since there is no mention here of the subject's power as the means for such causation, this means is best understood as “one's representations”. Given the temporal proximity of the composition of *MM* and *Anth.*, Kant is probably expressing the same idea in *Anth.* 7:251, such that “its effect” refers to the effect of the representation of something in the future. If this is correct, desiring

⁹³ In contemporary discussions on desire, there is a debate on whether all desires are for states of affairs. Kant's talk of desires for objects and actions might be taken to suggest that he would deny that all desires are for states of affairs. In speaking of desires for objects and actions, however, Kant might be speaking loosely, and so he might accept that these desires are “disguised” desires for states of affairs (e.g., my desire for Brandy is actually a desire that I would be drinking Brandy in the near future; and my desire for gambling is a desire that I would be gambling in the near future).

ice-cream is done by means of representing ice cream or having it in the future as the effect of one's representation of ice cream (or of having it).⁹⁴

The idea that desires involve representations of *future* objects gains support from another passage in *Anth.*, where Kant says "Every desiring [*Begehren*] contains a (doubtful or certain) foresight (*Voraussehen*) of what is possible through it" (*Anth.* 7:185); and he explains the faculty of foresight (*Vorhersehungsvermögen*), as "the faculty of visualizing something as taking place in the future" (*Anth.* 7:182). Since every desiring involves foresight, i.e., visualizing some future object or state of affairs, desires are probably directed at bringing about these future objects or states of affairs (by means of their representations). And the idea that desires are future-directed in this way appears in several other passages, from different junctures in Kant's career: as early as 1764, Kant says that even if we had never defined what desire was, we could say with certainty that "every desire [*Begierde*] presupposed the representation of the desired [*des Begehrten*]; that this representation was a foresight [*Vorhersehung*] of what was to come in the future; that the feeling of pleasure was connected

⁹⁴ Note that in representing the object as caused by its representation I need not *believe* that the representation can cause the object (let alone that it suffices for causing it). It might be thought irrational to represent an object as caused by its representation without believing that the representation can cause it. But Kant's view appears to be that we are susceptible to such irrational acts of representation, committed whenever we wish for something. For he characterizes wish as an act of the faculty of desire that is not joined with consciousness of the ability to bring about the desired object (*MM* 6:213). In lectures, Kant's names for desires that are not wishes are "active" (*LA* 25:1335), "inciting" (*LA* 25:577) or "practical desires" (*LA* 25:796; *LM* 29:1013). He says they are "a basis of the effort of our power to make something actual" (*LA* 25:1335) and that they are "the basis of determination, to make the object actual for activity, and to procure." (*LA* 25:577). Although in *Anth.* Kant characterizes wish as desire without exercising power to produce the object, in other places he speaks of it as a desiring without awareness that the agent can bring about the desired object (*MM* 6:213); with awareness that one cannot bring about the desired object (*LA* 25:795; *LM* 29:1013-4); and as desiring of an object which is not within our control (*LM* 29:895) (see Engstrom 2010, 46, for a similar point). These characterizations are compatible and may be combined with the one in *Anth.* 7:251, however, such that in wishing we do not exercise our power to produce the object because we believe that we cannot produce it (and this is actually the case), or do not believe that we can produce it.

with it; and so forth" (*Inquiry* 2:284); in a draft note from the 1770's, he says that desire is "The feeling in view of the future (condition)" (15:733, translation mine); in his *LR*, which were given around the mid 1780's, Kant is recorded saying that "every desire is directed only to something possible and future" (*LR* 28:1059); and in a lecture from 1782-3, that "All desires are always directed to something in the future." (*LM* 29:894). So, there is strong evidence that desires involve representations of future objects (in the abovementioned broad sense); and given that the faculty of desire is the faculty to cause objects by means of their representations, it is reasonable that desires are directed at causing objects in the future (relative to the time when these objects are desired) by means of their representations.⁹⁵

In other passages, however, Kant acknowledges that some desires are directed at past objects. He says, for example, that we can desire to "make what has happened not have happened" (*CPJ* 5:178 fn.); and he also mentions this desire in a lecture, saying that "we often wish that something would not have happened which, however, now is impossible." (*LA* 25:1335). These past-directed desires, however, are *wishes*, that is, desiring without exercising power to produce the object (*Anth.* 7:250), and so we can reconcile the conflicting evidence by reading the claims about the future-directedness of desires as referring to *active*, non-wishful desires alone, that involve exercising power to produce their objects. If this is correct, desires - except for certain wishes - are future-directed.

3.4. Desires as Dispositions

I have just argued that "desire" in the first sentence of *Anth.* 7:251 refers to both intellectual and sensuous desires, and that desires usually involve representing something in the future as caused

⁹⁵ One option is that desire involves actual pleasure in a representation of a future object; another is that it involves an expectation of pleasure in such a representation. I will consider these two options in 3.6.

by this representation. We can now proceed to examine the trickier part of Kant's characterization of desire, the “self-determination of the subject's power” component. Here too, it would be helpful to distinguish between the “self-determination” and the “subject's power” component.

Starting with the latter, while it is not immediately clear what “the subject's power” refers to, one reasonable candidate is the faculty of desire itself. For if desires are determinations of the faculty of desire, it is natural to read Kant's characterization of desire as “a self-determination of the subject's power” as referring to a self-determination of the faculty of desire (otherwise desires would be determinations of the faculty of desire, and self-determinations of another power). Against this reading one might object that the power that is self-determined in desire must be one whose exercise is sufficient for causing objects, and the faculty of desire is insufficient for this purpose, as Kant says that our mechanical powers “have to be determined through that representation (of the desired object - U.E.) in order to realize the object” (*CPJ* 5:178). My response to this objection is that exercising the faculty of desire may involve determining one's mechanical powers to action, and if this is the case, self-determination of the faculty of desire could suffice for “realizing” objects. If, however, exercising the faculty of desire does not involve determining our mechanical powers to action, then exercising it would be insufficient for producing desired objects, and so “the subject's power” should not be understood as referring to it. An alternative would be to read it as referring to *the subject's power as whole*, i.e., the sum of the subject's powers sufficient for producing objects (including, of course, mechanical powers). So, it is either the faculty of desire, or the subject's power as whole that is self-determined in desire.

Before I proceed to examine what self-determination means, I want to consider Morrisson's alternative reading of "the subject's power", on which it is *the power of imagination* that is self-determined in desire. Morrisson interprets *Anth. 7:251* as saying that a desire is "a determination of the power to imagine something in existence as a result of our imagining it"; to have a desire in this sense, "is just to imagine something in existence as a result of imagining it." (Morrisson 2008, 29). He concedes that it is not immediately clear how this understanding of desire is linked to the idea that desire is "that by which something is brought into existence via the idea that we have of this thing" (*ibid.*), which he takes to be implied by Kant's characterization of the faculty of desire in *MM 6:211*; and he attempts to explain how the two characterizations may be combined. There are other reasons, however - both philosophical and textual - to think this reading of 7:251 is incorrect.

Starting with the former, on Morrisson's reading, in desiring we imagine something in existence as a result of imagining it. It is not clear, however, what the difference between these two mental states is supposed to be, nor that there is one. And even if these mental states were to differ, it would be odd for Kant to call the act of imagining something in existence as a result of imagining it "desire". Suppose, for instance, that I imagine that it is actually raining outside ("imagining something in existence") as a result of imagining rain ("imagining something"). It seems that I can do all this without thereby desiring that it would be raining, and even while desiring that it would not be raining. As far as textual considerations go, Morrisson's reading relies on his modified translation of the above passage: "Desire (*appetitus*) is the self-determination of the power of a subject to imagine something in the future as an effect of such imagination" (Morrisson 2008, 29).⁹⁶ This translation is misleading, however, insofar as it

⁹⁶ The original text reads: "Begierde (*appetitus*) ist die Selbstbestimmung der Kraft eines Subjects durch die Vorstellung von etwas Künftigem als einer Wirkung derselben." (*Anth. 7:251*).

gratuitously identifies the power of the subject that is self-determined in desire with the power to imagine something in the future. For although the text may be read such that this self-determination of the power of the subject occurs *through (durch)* the imagining of something in the future, it does not say that it is a self-determination *of* the power to imagine something in the future. I therefore submit that the reading on which desire is self-determination of the power of imagination is not a good alternative to the two I've proposed (i.e., self-determination of the faculty of desire, or of the subject's power as a whole).

This leaves us with the mysterious “self-determination” component of desire. “Determination” is a technical term with several uses in Kant, but in its most general sense, it means a positing of a predicate to the exclusion of its opposite (A591/B599; 1:391). In the context of a desire, it would therefore consist in a certain predicate being true of the desiring subject or of its power. And since desire involves representing something in the future as the effect of its representation, and since the faculty of desire can cause objects by means of their representations, the “predicate” that is true of the desiring subject should be one whose “positing” *can* cause objects by means of their representations. I therefore propose that the self-determination involved in desire be understood as a motivation or disposition⁹⁷ to act in order to bring about the desired object. On this proposal, desires *can* lead to actions - understood as efforts to bring about objects by means of their representations - but they need not do so.

⁹⁷ “Disposition” is sometimes used to translate Kant's terms *Gesinnung*. But although on my suggestion a desire is a disposition, it is not a *Gesinnung*, which is “the first subjective ground of the adoption of the maxims”, and “can only be a single one, and it applies to the entire use of freedom universally.” (*Rel.* 6:25). While an agent has many inclinations, she can only have one *Gesinnung*, understood by Allison as “the enduring character or disposition of an agent, which underlies and is reflected in particular choices.” (Allison 1990, 136).

In contrast to my reading, Frierson takes desires to be much more closely connected to actions, arguing that desires are causes of action in the absence of external impediments. It is not immediately clear what kind of impediments could prevent desire from leading to action, according to Frierson, but one such impediment is what Frierson calls “subjective incapacity”, such as the inability to climb a tree that might prevent one from getting a desired mango (Frierson 2014b, 56). Another reasonable candidate for an impediment that could prevent desire from causing action would be some event outside the agent's control at the time of desiring (as when my desire to kick the ball fails to cause the action because I am tackled by another player).

It might be thought that these impediments need not be external to the agent or outside her voluntary control - they can be external merely to the potentially action-causing desire itself. If this is so, then external impediments also include the self-control exercised when resisting the temptation to act on a desire; and perhaps an external impediment could also be another, stronger desire that may prevent the potentially action-causing desire from bringing action about. But this liberal understanding of “external impediment” conflicts with the contrast Frierson draws between Kant’s notion of desire and the customary English sense of the term, “whereby one can desire something without actually pursuing it.” (Frierson 2014b, 56). The natural implicature here is that one *cannot* desire something in Kant's sense without actually pursuing it. And if this is the case, then the impediments that could prevent desire from causing action cannot be other desires and self-control; for if these impediments could prevent action, one *could* desire something in Kant's sense without actually pursuing it. I therefore think Frierson is better understood such that the impediments that may prevent desire from causing action are restricted to events and states of affairs outside the agent's voluntary control at the time of desiring

(including “subjective incapacities”). It is in the absence of *these* impediments that Frierson thinks desire necessarily causes action.

There are two problems, however, with Frierson's understanding of desire as the cause of action in the absence of external impediments: (1) it conflicts with ordinary language use of “desire”; and (2) some mental states associated by Kant with the faculty of desire need not cause action in the absence of external impediments. Frierson addresses these problems in the following passage:

Kant's notion of desire is more closely connected to choice and action than the customary English sense of desire, whereby one can desire something without actually pursuing it. Once one has a desire in this general Kantian sense, one is committed to action, and action follows necessarily in the absence of hindrances... In contrast to typical English usage, for Kant desires mark an end to deliberation, not factors taken into account in deliberation. Thus a Kantian 'desire' (*Begehren/Begierde*) might better be called a 'volition' (as Kant does, identifying it with *Wollen* at 25:1334). Kant develops other categories – such as 'wish' (*Wunsch*), 'inclination' (*Neigung*), and 'ground of desire' – that serve the purpose of what we might call 'mere' desires.' (Frierson 2014b, 56).

Frierson therefore does not think that the incongruence between ordinary English and Kant's notion of “desire”, as he understands it, is a decisive reason to reject his reading. And since he contrasts “a Kantian desire” with “other categories” such as wishes and inclinations, and classifies the latter as “grounds of desire” (Frierson 2014b, 56, fn.12; cf. 83), his view must be that the latter are not desires in Kant's proper sense. In other words, while desire in Kant's proper sense is a volition, the outcome of practical deliberation that causes action in the absence of external impediments, wishes and inclinations are “mere desires”, or “grounds of desire”; and Frierson proceeds to argue that grounds of desire explain how a sensory cognition of an object causes a proper desire (a “volition”) for it (Frierson 2014b, 69).

What Frierson's reading gets right, I think, is that *some* of the mental states Kant associates with the faculty of desire actually bring about action. Kant's name for desires that do that, as Frierson himself notes, is “volitions”. In the passage Frierson refers to, Kant says that desire “often produces an effort aimed at the existence of the thing and is thus the cause of an action insofar as the thing is in my control... Desiring [*Das Begehren*] is also called volition.” (*LA* 25:1334). Commenting on this passage, Frierson says that Kant identifies *desire* with volition; but in fact, Kant identifies *desiring* with volition. This is important because it means that this passage need not be taken as evidence that desire as such is the cause of action insofar as it is under one's control; it may rather be read such that there is a certain kind of desire, namely, volition (or “desiring”), which causes action insofar as it is under one's control.⁹⁸ Frierson's “action-based” account, I propose, holds for desires of this kinds, that is, for *volitions*.⁹⁹

Where Frierson goes astray, however, is in reading Kant's characterizations of the faculty of desire as implying or suggesting that desire as such - that is, all desires - causes action in the absence of external impediments. The first passage he cites to support this claim is one where Kant says that desire is that in the thinking being “which [corresponds to] the motive force in the

⁹⁸ (1) “Desiring” as used here is thus not any mental act that involves having a desire (the sense I have argued it has in *Anth.* 7:251), but rather one that actually leads to action. (2) Kant also says in *LA* 25:1334 that desire *often* produces an effort aimed at the object's existence, which suggests that desire need not cause action. Frierson's rejoinder might be that Kant's qualification is meant to take care of cases where external impediments prevent desire from causing action. But earlier, Kant speaks of inclination as habitual sensible desire (*LA* 25:1334), and a bit later about idle desires (*LA* 25:1335). Assuming these mental states need not cause action (as Frierson concedes), Kant should not have called them “desires” if his view were that desires are causes of action insofar as it is under our control.

⁹⁹ It might be objected that although volition *produces* or *leads to* action, it cannot be said to *cause* it because causation entails necessitation, which is incompatible with freedom. But Kant speaks of desire (in its guise as volition, if my reading is correct) as “the cause of an action insofar as the thing is in my control.” (*LA* 25:1334); and he speaks of the faculty of desire as the faculty to cause objects (*MM* 6:212) and of the will as a kind of causality (*GMM* 4:446). So, Kant has no problem using causal language in referring to human action or volition. And this may be because although volition (as an act of the faculty of desire) *necessitates* action, it is free in the sense that it is not *necessitated* by anything.

physical world... [L]iving things do something according to the faculty of desire, and lifeless beings do something when they are impelled by an outside force." (LA 25:577).¹⁰⁰ But Kant need not be read as saying here that all desires are causes of action; he may rather be speaking of the *function* of desire, which is analogous to that of a physical motive force: it enables living things to "do something", to act. But while the function of desire is to produce action, not all desires fulfill this function.

Similarly, in the second passage Frierson refers to, Kant characterizes the faculty of desire as "*the faculty of the soul for becoming cause of the actuality of the object through the representation of the object itself*" (LM 29:1012). Frierson comments on this passage, saying that desire is defined as "a representation that leads to action, that 'becomes cause of the actuality of an object'." (Frierson 2014b, 55). But Kant's point here may be that it is the function of the faculty of desire (and of desire) to bring about action, and from this it does not follow that every desire actually fulfills this function.¹⁰¹ Since the view that desire leads to action (in the absence external impediments) requires denying that inclinations and wishes are desires in Kant's proper sense, we had better adopt the weaker reading, on which it is the function of desire to bring action about, such that desires are motivations or dispositions to action that need not produce it. And this conception of desire applies both to volitions (i.e., desires that actually lead to action and motivate the actions they bring about), and to wishes and inclinations (desires that need not lead to action, and whose motivating force could remain a mere potential).

¹⁰⁰ I am using the translation used by Frierson for the purpose of engaging with his argument.

¹⁰¹ An analogy might help to clarify this point. Suppose that it is the function of the human being to engage in rational activity of some sort. While this sets a normative standard by which we can assess human behavior, it does not mean that all human activity is rational. Similarly, if the function of desire is to cause action, this provides a standard for assessing desires (though not the only standard, for some desires should not lead to action), but does not mean that all desires meet this standard.

3.5. Desire and Pleasure

If what I have argued is correct, desires are dispositions or motivations to action that need not bring it about, and are normally (with the exception of certain wishes) directed at future objects. While this takes care of the characterization of desire in *Anth.* 7:251 and of the faculty of desire (in *MM* 6:212 and *CPrR* 5:9 fn.), in other passages we find a different characterization, on which desires are connected with pleasure or satisfaction, and even identified with them. These passages, which seem to lend support to the “feeling-based” account, give rise to two important interpretive questions: (1) is desire a kind of pleasure, necessarily connected with it, or both? (2) if desire is connected with pleasure but itself is not a pleasure, how does it differ from the pleasure it is connected with? To anticipate, I will argue that desire is necessarily connected with pleasure, or contains pleasure as its component, but itself is not a pleasure; and that desire differs from the pleasure it is connected to because desires are (normally) future-directed, which allows them to motivate action immediately, while pleasures are not future-directed, and so they must cause desires in order to produce action.

First, consider the evidence that all desires are *connected* with pleasure or satisfaction. In *MM*, Kant says that "In the case of desire or aversion, first, *pleasure* or *displeasure*, susceptibility to which is called *feeling*, is always involved;¹⁰² but the converse does not always hold, since there can be a pleasure that is not connected with any desire for an object but is already connected with a mere representation that one forms of an object (regardless of whether the object of the representation exists or not)." (*MM* 6:211). Kant's point here is that certain pleasures - later called “merely contemplative pleasures” or “inactive satisfactions” (*unthätiges Wohlgefallen*, *MM* 6:212) - are in the mere representations of their objects and so may be

¹⁰² I am using Grenberg's translation of the part of this sentence up to the semicolon (Grenberg 2001, 160, fn. 18), for the reason specified in fn. 16, Chapter 1.

disconnected from desires,¹⁰³ while desires are always connected with pleasures. And as we have seen earlier, there are two kinds of desire, according to Kant: those caused by pleasures (“desire in the narrow sense”), and those that cause pleasures (called “sense-free inclination” when habitual). The pleasure that is not necessarily connected to desire is “not at bottom a pleasure in the existence of the object” (*MM* 6:212), and so it is reasonable to infer that the pleasure that is necessarily connected to desire is (“at bottom”) in the object's existence (rather than in its mere representation). But what's important for our current purpose is that pleasure - even in its guise as practical pleasure which necessarily involves desire - is *distinct* from desire.

Similarly, in a lecture, Kant asserts that desiring is connected with *satisfaction*, when he explicates the concept of desiring (*Begehren*) as “*the representation of the object, which is connected with satisfaction in its actuality, and which is the ground of producing it.*” (*LM* 29:1013). According to this passage, desiring is the representation of an object that is a ground - a reason, or a cause - for producing the object, and this representation is connected with satisfaction in the actuality of the represented object. In the previous chapter we have seen that satisfaction is Kant's term for all feelings, be they sensuous or intellectual, and so this passage suggests that desire is *connected* with a feeling (of pleasure, or satisfaction) in an object's existence or actuality, but not itself a feeling. Call this reading, on which desire is connected to a feeling of pleasure or satisfaction, but itself is not a feeling, (a). Combined with our previous

¹⁰³ It is not clear whether Kant's view is that pleasures in representations are necessarily disconnected from desires, or whether they are *not necessarily connected* with desires (such that some are connected with desires). The latter is suggested when Kant says “that pleasure which is not necessarily connected with desire for an object, and so is not at bottom a pleasure in the existence of the object of a representation but is attached only to the representation by itself, can be called merely contemplative pleasure or *inactive satisfaction* [*Wohlgefallen*].” (*MM* 6:212). If pleasure in a representation may be connected with desire, there is room to distinguish between contemplative pleasures that are connected with desires, and contemplative pleasures that are not connected with desires (i.e., “disinterested pleasures”).

characterization of desire as a disposition, the result is this: desire is a disposition to action, normally directed at producing a future object, connected to a feeling of pleasure or satisfaction in an object's existence.

In contrast, in other passages, Kant's view appears to be that desire *is identical with* a kind of satisfaction, namely, satisfaction in an object's existence. For example, Kant says that "Satisfaction with the actuality of the object is desire" (*LA* 25:577), and that "Desire is the satisfaction [*Wohlgefallen*] in the existence of the object." (*LA* 25:1334). That desire need not be an intellectual feeling is suggested when Kant says elsewhere that desire is a *pleasure*, and specifically "a pleasure insofar as it is a ground of an activity for determining certain representations of the object" (*LM* 28:254); and that desires, also in their guise as volitions, are satisfactions in objects' existence, is strongly suggested when Kant says that "to will something and to have satisfaction in its existence... are identical" (*CPJ* 5:209), which suggests that volition, the act of the faculty of desire that normally leads to action, is satisfaction in an object's existence.¹⁰⁴ Moreover, just before he says in the passage cited above (*LM* 29:1013), that desire is *connected* with satisfaction, Kant identifies desire with satisfaction, saying that it is "satisfaction with respect to the actuality of the object <*complacentia respectiva actualitatis objecti*>" (*ibid.*). These passages suggest that desire - both as a possible ground of volition and as volition - is a feeling in response to an object's existence or actuality. And if this is the case, then either: (b) desire is a kind of feeling, but is not causally connected with feeling; or (c) desire is

¹⁰⁴ It might be thought that we can make sense of the mixed evidence by reading passages on which desire is a pleasure or satisfaction as referring to desires as possible grounds of action, and passages on which desire is connected with pleasure as referring to desires as volitions, i.e., actual grounds of action. But the fact that Kant says that willing is satisfaction in an object's existence (*CPJ* 5:209) suggests that actual grounds for action are satisfactions; and if inclination is a desire in the narrow sense (*MM* 6:212), it is caused by pleasure, which means that some possible grounds of action (namely, inclinations) are connected with pleasure.

both a kind of feeling and causally connected with feeling. What is attractive about these readings, I think, is that they allow us to regard the mental states of the faculty of feeling and those of the faculty of desire as two species of one genus (i.e., “feeling”, or “pleasure” / “satisfaction”); and this suggests that although the mental states we ordinarily refer to as emotions come from two different mental faculties in Kant, they may nonetheless be explained by a single theory, as there is one Kantian category that applies to all of them.

But despite this merit, there are also good reasons to reject these readings. First and foremost, the idea that desire is a kind of feeling (of pleasure, or satisfaction) threatens to collapse Kant's feeling-desire distinction. For recall that Kant's name for the faculty of feeling is "the feeling of pleasure and displeasure" or "the feeling of satisfaction and dissatisfaction". This suggests that all feelings, pleasures and satisfactions, are products of the faculty of feeling, and if this is the case, desires cannot be feelings, satisfactions or pleasures, as this would mean they are products of the faculty of feeling, not of the faculty of desire. In defense of (b) and (c), it might be argued that although all the mental states of the faculty of feeling are feelings of pleasure or satisfaction, this does not entail that all feelings, pleasures and satisfactions are products of this faculty; and given that Kant refers to desire as a satisfaction (*LA* 25:577) a pleasure (*LM* 28:254) and a feeling (*LA* 25:795), it is possible that satisfaction (or pleasure, or feeling) in mere representation would be a product of the faculty of feeling, while satisfaction (or pleasure, or feeling) in an object's existence would be a product of the faculty of desire.

There are other problems with the idea that desires are pleasures or satisfactions, however. First, it conflicts with passages that suggest that desire is distinct from pleasure or satisfaction (i.e., *MM* 6:212, and *LM* 29:1013, respectively). Second, in the passages where Kant characterizes desire as self-determination (*Anth.* 7:251) and the faculty of desire as the faculty to

cause objects by means of their representations (*MM* 6:211; *CPrR* 5:9 fn.), there is no mention of pleasure as essential to desire; these passages rather suggest that it is the ability to produce action that is essential to desire. I therefore submit that both (b) (desire is a kind of feeling, not causally connected to feeling) and (c) (desire is both a feeling and causally connected to feeling) are implausible.¹⁰⁵

This leaves us with the option that desire is connected to feeling, but itself is not a feeling (a). In contrast with (b) and (c), this reading does not threaten to collapse the feeling-desire distinction. And although it is in tension with passages where Kant speaks of desire as a satisfaction or pleasure in an object's existence, these passages may be read as referring not to desire itself, but rather to a phenomenon that is necessarily connected with it, or to a necessary component of it, such that whenever we take satisfaction in an object's existence we desire the object. If this reading is endorsed, these passages are not saying that desire is identical or consists in a satisfaction, and so (a) remains a viable option.

But if desire is not a pleasure in an object's existence, but merely necessarily connected with it or contains it as a component, how do the two differ? And more specifically, if, as I have argued in the previous chapter, pleasure is a behavioral disposition, and, as I have argued earlier in this chapter, desire is a disposition to action, why is it that the former cannot but the latter can

¹⁰⁵ In defense of (c), it might be argued that it deserves further consideration because it has the additional merit that it incorporates both passages where desire is identified with satisfaction and those on which it is connected with pleasure. But despite this merit, (c) is vulnerable to another problem, namely, that the pleasure which desire is connected with appears to be the same as the one it is identified with, i.e., pleasure in the object's existence. That the pleasure connected with desire is in the object's existence is suggested when Kant contrasts practical pleasure with inactive delight, saying of the latter that is "not at bottom pleasure in the existence of the object" (*MM* 6:212). The natural implicature is that practical pleasure *is* ("at bottom") in the existence of the object. And if this is the case, (c) amounts to the implausible view that desire is satisfaction in an object's existence caused by satisfaction in an object's existence.

motivate action immediately? Grenberg addresses these questions when she says that "Technically... there is a distinction to be made between feeling and desire: feeling, an element of the agent's sensible nature, influences or, in Kant's language, 'determines' (*bestimmt*) the faculty of desire 'subjectively' so as to produce what Kant calls an inclination [*Neigung*] and what we commonly call a 'desire'. For the purposes of describing action, there is, however, little distinction to be made between the possession of a practical pleasure and that of a desire." (Grenberg 2001, 163). Grenberg's view is thus that feelings of pleasure and displeasure are "elements of the agent's sensible nature", and earlier they are explained as reactions to representations that have a certain degree of *strength* (Grenberg 2001, 160); these reactions to representations, which are determinations of the faculty of feeling, determine, in turn, the faculty of desire, and produce inclinations or desires.¹⁰⁶

Although Grenberg qualifies her claim such it is "for the purposes of describing action" that there is "little" distinction between desires and practical feelings, it is not clear what the distinction amounts to, on her account. A practical feeling, according to Grenberg, is "a feeling directed toward bringing about the existence of its object" (Grenberg 2001, 161); and feelings

¹⁰⁶ It might be thought that Grenberg's account applies only to a particular species of the mental states that originate in the faculty of desire, namely, to inclination. But Grenberg seems to think that for Kant, all subjective determinations of the faculty of desire (i.e., actual determinations, as opposed to possible rational determinations of this faculty in its guise as will) are inclinations (Grenberg 2001, 163, fn. 27). Accordingly, I take it that her account is meant to apply to all the mental states which originate in the faculty of desire. Although the idea that all subjective determinations are inclinations might seem implausible, I tend to think that this comes close to Kant's actual view, on which the vast majority of desires in human adults are inclinations, i.e., habitual desires, in the sense that they are directed at objects one is acquainted with. This much is suggested by the fact that when he presents the different "levels" of the faculty of desire (e.g., *Rel.* 6:29 fn., *Anth.* 7:265), Kant does not mention that there is an intermediate level between instinct, which precedes being acquainted with an object or having a concept of it, and inclination. If this is correct, inclinations need not be deeply ingrained in or particularly telling of their agent's character: any desire for an object of a kind one is acquainted with or has a concept of is an inclination ("Inclinations are desires for objects of an entire genus.", as Kant says in *LA* 25:1140-1). This reading has the advantage that it can account for the fact that when speaking of human action, Kant usually uses "inclination" (and rarely uses "desire").

exert a certain force on us (ibid.). So, practical feelings seem to have all the ingredients necessary for motivating action. However, Kant's view is that practical feelings cannot motivate action directly; they rather do so by producing desires (or, as Grenberg says, by “determining” the faculty of desire). So, there must be some difference between practical feelings and the desires they are causally connected with, and Grenberg neglects to say what this difference is. In her defense, it might be argued that practical feelings differ from desires in that the former are products of the faculty of feeling, while the latter are products of the faculty of desire. This is certainly true, but does not solve the problem, as it still remains unclear why products of the faculty of desire can but those of the faculty of feeling cannot motivate action immediately.¹⁰⁷

How do practical pleasures differ from the desires they are connected to, then? A possible answer emerges once we consider the temporal directedness of desires. Earlier we have seen that desires (expect for certain wishes) are directed at future objects. And in some passages, Kant contrasts the future-directedness of desires with the present-directedness of feelings. For example, in a lecture from around 1776, Kant is recorded as saying that "The desire is not a perception of what is actual, but merely of what is possible and future. Feeling however aims at the present." (*LA* 25:589). Here, desires and feelings are distinguished according to their temporal-directedness: desires aim or are directed at “what is possible and future”, feelings aim at the present. Approximately two years later, Kant is recorded expressing the same idea: "Of Desires. It is a feeling in regard to what is future. But the effect of the representation of what is

¹⁰⁷ Frierson says he follows Grenberg "in downplaying the distinction between pleasure and desire." (Frierson 2014b, 60). In contrast with Grenberg, however, he does not think pleasures are directed at producing objects; he rather contrasts desires, which are “objective” in the sense that they are directed at producing objects, and pleasures which are “wholly subjective”, both in the sense that they reflect something about the subject, and in the sense that they aim to preserve themselves as mental states (Frierson 2014b, 59).

future on our feeling brings us into activity. The feeling of the present incites no activity, but when it means that it will endure, then an activity will soon be incited." (*LA* 25:795). Here, Kant goes as far as identifying desire with feeling, but for the aforementioned reasons I think it is best to read these claims such that feeling is necessarily connected to desire, or is a necessary component of it. The important thing for our current purpose, however, is that desire is characterized here as future-directed, the effect of a representation of a future object, and contrasted with "the feeling of the present". And although these passages are from lecture notes, Kant draws a similar distinction between affects and desires in the published *Anth.* For after he characterizes desire as self-determination through the representation of *something in the future* (*Anth.* 7:251), he proceeds to say that affect is "the feeling of a pleasure or displeasure in the subject's *present* state" (*ibid.*, italics mine). So, at least in some passages, Kant does not identify desires with practical feelings; he rather distinguishes between desires and feelings, or desires and affects, in terms of their temporal directedness.

The most straightforward way of making sense of the idea that desires are future-directed while feelings are present-directed is that desires are normally in response to (or aspects of) representations of future objects, while feelings are normally in response to (or aspects of) representations of present ones. If this is correct, and if desire is not itself a feeling of pleasure, then my desire for coffee, for example, is in response to (or aspect of) a representation of having coffee in the very near future, but is causally connected with a feeling of pleasure in response to a coffee that I am perceiving at this moment. The problem with this reading, however, is that some of our desires seem to be caused by feelings of pleasure in future objects (as opposed to present ones), as when my desire for coffee is caused by pleasure at the thought of having coffee in a few minutes. Moreover, if there are such pleasures, then both they and the desires they cause

are future-directed, and so they cannot be distinguished on the basis of the temporality of their associated representations.

I therefore think we should adopt an alternative reading, on which feeling is present-directed not in the sense that it is in response to a representation of a present object, but rather in the sense that it *refers to the subject's state*, as it is affected by the object, and so it has no temporal directedness. If this is correct, then although my desire for coffee may be aroused by pleasure at the thought of having it in a minute, the feeling of pleasure itself is neither present nor future-directed, as it refers to my affective state, not to temporal objects; the associated desire for coffee, in contrast, *is* future-directed because it is a self-determination to get the coffee by means of a representation of a future object (the coffee, or my having it in a minute). And so, the reason that feelings must cause desires if they are to produce action is that in themselves, even when caused by representations of future objects, feelings lack the future-directedness necessary for immediate motivation. This, however, is not because feelings are present-directed, but rather because they have no temporal directedness: they are not about objects in time, but rather about the way we are affected by objects (independently of their temporality).¹⁰⁸

¹⁰⁸ In a *Reflexion* note passage, Kant says that "The faculty of desire differs from... the judging of satisfaction and dissatisfaction: a. because desire refers only to future time, satisfaction to all times; b. desire is related to one's own power, to produce the object of satisfaction; hence we regard every desire for that which is not within our power as absurd. It is that within us by which we determine our power according to a foreseen (*vorhergesehen*) pleasure or displeasure." (R. 1008 15:447, translation mine). If we read "satisfaction" in "a" as referring to satisfactions that are not desires, i.e., to feelings, the result seems to be that feelings may be directed at the future. But Kant's point may be that feelings are in response to representations of past, present and future objects, while desires are directed at future objects. Again, the point is not that desire is a future-directed feeling, but rather that feeling refers to the subject and so it has no temporal direction, while desire is directed at bringing about a future object. I will address the possibility, suggested in this passage, that the pleasure that desire involves is an expected pleasure in the next section.

3.6. Possible Objections

If what I have argued is correct, desires are dispositions to action normally directed at future objects, and so they differ from the practical feelings they are causally connected to, which refer to the way we are affected by an object; the future-directedness of desires, I've suggested, is what enables them to motivate action immediately (but of course, they need not lead to action, and so wishes and inclinations are desires in this sense too). On this reading, although they are different from practical pleasures, desires always involve such pleasures. Before concluding, I want to consider two possible objections to my suggested account.

Frist, one might object that if, as my account has it, desires (+) involve pleasure (+), they cannot motivate action. For as we have seen in the previous chapter, on one of Kant's characterizations, pleasure (+) is "The consciousness of the causality of a representation with respect to the state of the subject, **for maintaining** it in that state...in contrast to which displeasure is that representation that contains the ground for determining the state of the representations to their own opposite (hindering or getting rid of them)." (*CPJ* 5:220). Pleasure (+), according to this passage, is consciousness of a causality for maintaining one's state, that is, a ground for maintaining one's state, while displeasure is a ground for changing it. And if this is correct, and if desires (+) involve pleasures (+), they involve grounds for maintaining one's state. This, however, suggests that while aversions can motivate us to act in order to change our state, desires (+), by virtue of involving grounds for maintaining one's state, involve grounds for inaction, that is, for doing nothing. In other words, we are motivated by aversion, not by desire (+). And this consequence conflicts both with common sense and with ordinary language use of "desire" and "aversion" (for we say, for example, that it was my desire (+) to see you - rather than my aversion to not seeing you - that made me drive for three hours).

One way to respond here would be to argue that although the pleasure connected to desire is an actual pleasure, desire itself is an *expectation of pleasure*. And there is some evidence to support this suggestion. For instance, in a reflection note Kant says of desire: "It is that in us through which we determine our power in accordance with foreseen (*vorhergesehenen*) pleasure or displeasure." (15:447, translation mine). Similarly, in a lecture he is recorded characterizing the faculty of desire as "the faculty to determine one's powers according to a foreseen pleasure and displeasure". (29:45, translation mine). And in another lecture, from 1781-2, Kant says that "In the case of desire, it is not the perception of the actual and objective, but rather an anticipated-sensation (*Vorempfindung*) of the future. Feeling refers to the present." (*Mensch.* 302, translation mine). According to these passages, desire is an expectation of pleasure, not an actual one. And this can solve the problem of the alleged inertness of desire (+) because an expected pleasure (+) is not a ground for preserving one's state but merely the expectation of having such a ground, and such an expectation *can* motivate action. Combined with the idea that desire is connected with pleasure, the result is that desire is an expectation of pleasure, causally connected with an actual pleasure. To pick up on our coffee example, my desire (+) for coffee, qua disposition to get it, is the expectation that coffee would be pleasant, caused by an actual pleasure (+) at the thought of getting coffee.

Another possible response would be to bite the bullet on the seemingly absurd consequence and say that although it might *seem* implausible that actions are always motivated by aversion, never by desire (+), there is evidence that suggests that Kant endorses this view. In *Anth.*, in a discussion about enjoyment and pain, which are sensuous feelings of pleasure and displeasure, he argues that "*no enjoyment can immediately follow another*"; rather, between one and another pain must appear... Pain is the incentive of activity, and in this, above all, we feel our

life; without pain, lifelessness would set in." (*Anth.* 7:231). If we take Kant's point here to be that we are always motivated by displeasure (or more specifically by pain), and Kant thinks of desire and aversion as involving pleasure and displeasure, respectively, then he also thinks that we are always motivated by aversion, never by desire (+).¹⁰⁹

Finally, it should be noted that although in *CPJ* 5:220 Kant speaks as if all pleasures (+) are grounds for maintaining one's state, which suggests that they cannot motivate action, we have seen in the previous chapter that there is also evidence that some pleasures (+) are grounds for producing objects. For in a passage from *FI*, Kant says that "**Pleasure** is a **state** of the mind in which a representation is in agreement with itself, as a ground, either merely for preserving this state itself... or for producing its object." (*FI* 20:230-1). Since it seems reasonable that the pleasure (+) that serves as a ground for producing its object is the one involved in desire (+), such pleasure (+) is distinct from a pleasure (+) that is a ground for preserving one's state, and so there is no reason to think it is inert.

The second objection I wish to consider is that my account is implausible because it involves a mysterious process whereby a motivationally inert mental state (a feeling of practical pleasure) can cause a motivationally efficacious mental state (a desire). And Grenberg's reading seems to fare better in this respect because it does not involve this mysterious process: desires, on her view, are caused by future-directed practical feelings, and so there is no ex-nihilo motivation added when feelings cause desires. My response to this objection is that although

¹⁰⁹ The idea that action is motivated by aversion rather than by desire (+) should be distinguished from the idea, commonly attributed to Locke, that we are moved to action by uneasiness. For on Locke's view, desire itself is uneasiness, which is what allows it to motivate (Locke 1975, II XXI 32). On the view I have suggested might be Kant's, desire cannot motivate because it involves pleasure. If this is Kant's view, he would agree with Locke that we are always moved to action by "negative" affective states, but disagree insofar as he would deny that desire consists or involves such a state. Locke's view, however, is not vulnerable to the objection that it conflicts with common use of "desire" as a motivator of action.

practical feelings are inert in the sense that they cannot motivate actions directly, by themselves, this does not entail that they cannot cause another mental state which can cause or dispose to action. And the reason that practical feelings, despite not being future-directed, can cause dispositions to action, is that they are *interested* pleasures, i.e., pleasures in response to objects the existence of which matters to us. Once we take an interested pleasure in an object or state of affairs, we represent the object or state of affairs in the future as caused by our representation, and it is by means of this representation that we are disposed to act so as to produce the object or state of affairs. So, although practical feelings cannot motivate action directly because they are not future-directed, their interested nature allows them to cause desires, which can motivate action directly.¹¹⁰

3.7. A "Motivation First" Holistic Theory of Desire

On the account I have defended, Kantian desires involve both a hedonic aspect (a pleasure), and a motivational aspect. But rather than causing action in the absence of external impediments, desires operate by disposing their agents to action, and wishes and inclinations are desires in this sense too. Moreover, although they are causally connected with practical feelings of pleasure,

¹¹⁰ It might be thought that there is another, more basic problem with my account: if desires (+) involve grounds for maintaining one's state, then one cannot have a desire (+) for something that is merely future and has never been experienced, since desire (+) must always be for the persistence into the future of what is presently pleasant. But although the characterization of pleasure in *CPJ* 5:220 suggests that desire (+) involves a ground for preserving one's state, this does not mean that desire (+) must involve a pleasure (+) in a present object, or an object that had already been perceived; for what one wants to preserve in desire (+) is not a present object or state of affairs, but rather a present mental state (which often requires taking action to bring about the represented object). And a mental state may be aroused by imagined objects which are not present at the moment, and even by objects one had never perceived. For instance, my desire (+) for the elimination of racial injustice involves pleasure (+) in the representation of a world where there is no such injustice, and this pleasure is a ground for preserving itself (according to *CPJ* 5:220). But what I want to preserve is not a world where there is no racial injustice (which is impossible, assuming this is not the current state of affairs), but rather the pleasure in the representation of such a world. And insofar as this pleasure is not a pure aesthetic one, it will cause a desire to bring about this world.

desires are future-directed, and this is what sets them apart from feelings of pleasure, and gives them their immediate motivational force.

In addition to being textually well-grounded, I believe this account is also philosophically compelling in that it foreshadows contemporary holistic accounts of desire. According to such accounts, having a desire is not a matter of having a *single* feature which is often associated with desire (such as a disposition to take action or to experience pleasure), but rather of having several of the features commonly associated with desires. These features include dispositions to action, dispositions to take pleasure in it seeming that the desired object or state of affairs is obtained, dispositions to believe that the desired object or state of affairs is good, and dispositions to attend to reasons to have the desired object (Schroeder, 2020).

Holistic theories of desire come in two kinds: functionalist and interpretationist. On functionalist theories, a desire is an internal state-type, that plays enough of the causal roles that features associated with desires do (e.g., Lewis, 1972). On interpretationist theories, desires are states of the whole organism that exist in virtue of exhibiting certain behaviors suggested by the features associated with desires, interpreted in accordance with certain principles (e.g., Davidson, 1980). Kant's understanding of desire as a mental state causally connected with feelings and actions in certain ways, may be understood as a functionalist holistic theory of desire. On this theory, desire is a disposition to action that is necessarily connected to a pleasure in an object's existence. Granted, if what I've argued is correct, then the dispositional component - the "self-determination of the subject's power" - is the essential component, and so unlike holistic accounts, Kant's account privileges one component over the others. But this disposition is necessarily connected to pleasure, and cannot occur without it (even in the case of rational desire, which necessarily causes pleasure). Moreover, arguably, because it always involves

representing a future object or state of affairs, desire may also be said to involve an evaluative component, whereby a represented future object or state of affairs is represented as good in some sense (which explains why the representation provides a reason for action). Kant's account is therefore attractive because it incorporates hedonic and dispositional components, as well as the evaluative component implicit in the future representation that desires involve, in a manner that anticipates contemporary functionalist holistic theories of desire.

We now have accounts of the two kinds of mental states pertinent to understanding emotions in Kant, i.e., feelings and desires. Mental states of these two kinds are similar in that they both involve behavioral dispositions, and they are closely connected because every desire involves a feeling. However, feelings cannot motivate action immediately because they refer to the subject's affective state, while desires (other than certain wishes) can motivate action immediately because they are future-directed. We can now proceed to look at the treatments Kant prescribes for these two kinds of mental states.

Chapter 4. Which Emotions Should Kantians Cultivate (and Which Ones Should They Discipline)?

4.1. Should a Good Kantian Cultivate Her Emotions?

So far, I have examined the nature of the two kinds of mental states pertinent to understanding emotions in Kant, namely, feelings and desires. I now want to look at the kind of treatment Kant prescribes for these mental states. One subset of these mental states on which Kant's position is fairly clear is those that stem from our so-called “moral endowments”, discussed in the Introduction to the *Doctrine of Virtue*.¹¹¹ Kant argues there that these “moral endowments”, which include moral feeling, conscience, love of human beings and respect (*MM*, 6:399-404), are natural predispositions of the mind that enable us to respond affectively to concepts of duty; and he is explicit that we ought to cultivate the first two.¹¹² It is less obvious, however, how, according to Kant, we ought to treat feelings and desires that do not stem from these endowments, and in particular, those that do not necessarily facilitate moral conduct.

Commentators have argued that Kant prescribes the cultivation of natural sympathetic feelings (Baron 1995, 211-17; Fahmy 2009, 37-40);¹¹³ of feelings, desires, attitudes and

¹¹¹ I say that certain mental states *stem* from these endowments because Kant refers to some of these endowments as susceptibilities or aptitudes that enable us to experience certain mental states, rather than as occurrent mental states. He says of moral feeling, for example, that it is “the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent with or contrary to the law of duty.” (*MM* 6:399); and he characterizes love of human beings as an aptitude (*Fertigkeit*) of the inclination to beneficence in general (*MM* 6:402). It is reasonable, however, that the mental states that stem from these endowments go by the same names. For discussions of these moral endowments, see Geiger (2011), and Guyer (2010).

¹¹² That we ought to cultivate moral feeling is apparent in Kant's saying “Obligation with regard to moral feeling can be only to *cultivate* it and to strengthen it through wonder at its inscrutable source.” (*MM* 6:400); that we ought to cultivate conscience is apparent when he says: “The duty here is only to cultivate one's conscience, to sharpen one's attentiveness to the voice of the inner judge, and to use every means to obtain a hearing from it” (*MM* 6:401).

¹¹³ It is not obvious whether the natural sympathetic feelings mentioned in *MM* 6:456-7 are different from the moral endowment of love of human beings (or “love of one's neighbor”), discussed in *MM* 6:401-2. I will address this issue in 4.3

dispositions that facilitate moral action (Baxley 2010, Chapter 2); of emotional sensitivity (*Empfindsamkeit*) (Sherman 1997, 168-9, Fahmy 2009, 40-2); and that he takes both moral feelings and the entirety of our sensuous desires as subject to cultivation (Papish 2007, 136). Others, in contrast, are skeptical that Kant recommends cultivating emotions. Thomason (2017), for example, argues that with the possible exception of the four “moral endowments”, Kant does not believe we should cultivate emotions, nor emotional sensitivity. And Borges grants that we should cultivate some emotions in order to accomplish benevolent actions (Borges 2004, 154); but she maintains that on Kant's view, the extirpation of the emotions is an ideal human beings should strive for, although they rarely achieve it (Borges 2008, 47; 2019, 107).

This disparity of interpretative positions suggests that commentators have different things in mind when speaking of “emotional cultivation” in Kant. This impression is reinforced once we note that it is not obvious what “cultivation” means in Kant; that he has no German term equivalent to “emotions” as used in contemporary philosophy; and that commentators use “emotions” to refer to different sets of mental states in Kant. Rather than resolving the apparent disagreement by sorting out the different senses scholars assign to “emotions” and “cultivation”, this chapter aims at moving the discussion forward by sticking to Kant's vocabulary for the mental states we call “emotions” (which, on our assumption, includes both “feelings” and “desires”), and to his terms for their proper treatments. My clue for developing the pertinent terminology comes from a passage where Kant says that “The sensations of human beings require cultivation and the inclinations require discipline.” (*LA* 25:622). This suggests that contrary to the tendency to speak in a wholesale manner about the appropriate treatment of emotions that do not stem from our moral endowments (mostly among those who are skeptical that Kant approves of emotions), different emotions require different treatments. And I shall

argue that this is Kant's actual view: feelings (except affects) require cultivation, while inclinations – although they can and perhaps may be cultivated – generally require discipline.

I begin with some background on the discipline-cultivation distinction in Kant, arguing that discipline is constraint by rules, and cultivation is the acquisition or improvement of a faculty or a skill for attaining rational but not necessarily moral ends. We morally discipline inclinations by scrutinizing our actions and performing our duty in the face of opposing inclinations. This process has the effect of cultivating moral feeling so that we are habituated into taking pleasure in the thought of dutiful action (4.2). I proceed with evidence that feelings (except affects) are to be cultivated, while inclinations generally require discipline; and I propose that we are to cultivate compassionate feelings by seeking out and not avoiding occasions for having them so that we may use them for performing morally required actions (4.3). Kant's prescription that we cultivate feelings and discipline inclinations makes sense because discipline involves preventing a mental state from motivating a forbidden action (or from motivating the omission of a required action), and inclinations can but feelings cannot motivate action immediately (4.4). Finally, although I read Kant as prescribing that we cultivate feelings, I argue that my reading is not vulnerable to Thomason's recent objections to emotional cultivation (4.5).

4.2. Discipline and Cultivation

Kant's most elaborate account of cultivation and discipline appears in *LP*.¹¹⁴ There, they are presented as two of the three components of education, but I will use this account to develop a broader notion of cultivation and discipline. The first component of education, care, is "the

¹¹⁴ For helpful discussions, see Loudon (2002, 38-44) and Munzel (1999, 279-88).

precaution of the parents that children not make any harmful use of their powers." (LP 9:441).

The second component of education is discipline:

Discipline prevents the human being from deviating by means of his animal impulses from his destiny: humanity ... Training is therefore merely negative, that is to say, it is the action by means of which man's tendency to savagery is taken away... Savagery is independence from laws. Through discipline the human being is submitted to the laws of humanity and is first made to feel their constraint. (LP 9:442)

Kant appears to be using "discipline" (*Disciplin*) interchangeably with "training" (*Zucht*) here, as he does in other places (LP 9:441; MM 6:485; CPJ 5:432). Accordingly, I take what he says of training to apply also to educational discipline: educational discipline is a process that takes away the human being's tendency to savagery. The tendency to savagery is a tendency to be independent of laws and to deviate from them on account of one's animal impulses. Educational discipline takes away this tendency by subjecting the human being to the laws of humanity and making him feel their constraint (*Zwang*).

I call this sort of discipline "educational" because Kant discusses it in the context of his theory of education, and it probably applies only to discipline of children by educational figures. But although the nature of its laws varies, "discipline" is characterized in several places in terms of constraint by means of laws or rules, or constraint that takes away a tendency to stray from rules. In *LA*, Kant says that "Discipline is the constraint (*Zwang*) of inclination in accordance with rules"; and that "under 'discipline' one understands the limitation of inclination through a certain rule." (*LA* 25:651, 1530). Kant is saying in these passages that discipline is the constraint or limitation of inclinations by means of rules. Such constraint, it seems, could be educational (if aimed at children's inclinations) or non-educational "self-discipline" (if an adult disciplines her

own inclinations). Call this kind of discipline, which consists in constraint of inclinations by rules, “practical discipline”.

The idea that discipline involves constraint appears also in the first *Critique*: "The **constraint** [*Zwang*] through which the constant propensity to stray from certain rules is limited and finally eradicated is called **discipline**." (A709/B737). The discipline Kant is talking about here is theoretical rather than practical, and so it is directed against a propensity to stray from theoretical rules. But the explanation of discipline as constraining a propensity (*Hang*) to stray from rules is similar to the one in *LP* (taking away a tendency to stray from laws). And so, Kant appears to have a generic notion of discipline that includes practical and theoretical discipline. “Discipline” in this generic sense is constraint by means of rules or laws that aims to take away a propensity to stray from them. And “practical discipline” consists in constraining inclinations by means of practical rules. Kant also introduces a distinction within practical discipline between moral and prudential discipline: the former is rational constraint of the propensity to stray from moral laws, the latter is rational constraint of the propensity to stray from rules of prudence (*LE* 27:360).

Cultivation (or “culture”, *Cultur*), the third component of education, is explained as "the procurement of skillfulness (*Geschicklichkeit*)", which is "the possession of a faculty (*Vermögen*) which is sufficient for the carrying out of whatever purpose." (*LP* 9:449). Educational cultivation is here understood as a process of acquiring a faculty that would enable us to perform certain actions and attain certain ends. Moreover, the characterization of cultivation as acquiring an ability is not limited to educational cultivation. In the first *Critique*, Kant gives a similar account, contrasting cultivation with discipline, and saying that the former "would merely produce a **skill**

(*Fertigkeit*)¹¹⁵ without first cancelling out another one that is already present." (A709/B737). In *MM*, he adds that natural perfection is a duty and consists in "the *cultivation* of any *faculties* [*Vermögen*] whatever for furthering ends set forth by reason" (*MM* 6:391); and he adds that there is a law for maxims of actions that says "Cultivate your powers of mind and body so that they are fit to realize any ends you might encounter" (*MM* 6:392). Kant therefore has a generic notion of cultivation, understood as acquiring or improving a faculty, a power, or skill for furthering rational ends, be they moral or non-moral; and since we have a duty to acquire or improve faculties, skills and powers for attaining rational ends, we have a duty to cultivate feelings, desires, or the capacity for having them, if they constitute faculties, powers, or skills for attaining these ends.¹¹⁶

There is, however, a broader sense of "cultivation" in Kant. Cultivation in this sense is not opposed to discipline, but rather applies to both discipline and cultivation in the narrow sense (acquiring a faculty for rational ends). This sense appears in *CPJ*, when Kant speaks of "cultivation of skill" (*Cultur der Geschicklichkeit*), and contrasts it with a "cultivation of training (discipline)" (*Cultur der Zucht, Disciplin*, 5:431-2). Here, cultivation can be either positive, "the foremost subjective condition of aptitude for the promotion of ends in general" (cultivation of skill), or negative, "liberation of the will from the despotism of desires" (cultivation of training)

¹¹⁵ *Fertigkeit* is sometimes translated as "aptitude" (for example, in the Cambridge translation of *MM* 6:402, 407).

¹¹⁶ It might be thought that we ought not to cultivate feelings because they are not faculties or powers, but are rather passive susceptibilities. But as noted earlier, Kant sometimes lists the feeling of pleasure and displeasure as one of the three primary mental faculties (*CPJ* 5:198 *FI* 20:245), which suggests that feelings, in their guise as enabling us to experience certain mental states, *are* faculties. Moreover, even if some feelings are not faculties or powers, they may be aptitudes or skills (*Fertigkeit*), in which case it is possible that they ought to be cultivated.

(*CPJ* 5:432). So “cultivation” seems to have a broad sense, which includes both discipline (“cultivation of training”) and cultivation in the narrow sense (“cultivation of skill”). In what follows, however, I will be using “cultivation” in the narrow sense, unless indicated otherwise.

This gives us some idea about what cultivation and discipline mean in Kant. But how do we discipline and cultivate ourselves and our feelings or desires? A complete answer to this question will emerge in the coming sections. In a passage in *LE*, however, Kant gives us some important clues:

...self-mastery rests on the strength of the moral feeling. We may have good command of ourselves if we weaken the opposing forces. But this we do when we divide them; hence we first have to discipline ourselves, i.e., to root out, in regard to ourselves, by repeated actions, the propensity [*Hang*] that arises from the sensuous incentive [*sinnlichen Triebfeder*]. He who would discipline himself morally must pay great attention to himself, and often give an account of his actions before the inner judge, since then, by long practice, he will have given strength to the moral motive [*Bewegungsgrund*], and acquired, by cultivation, a habit of displaying pleasure or displeasure [*Lust oder Unlust zu bezeigen*] in regard to moral good or evil. By this the moral feeling will be cultivated, and then morality will have strength and motivation; by means of this incentive [*durch diese Triebfeder*], sensibility will be weakened and overcome, and in this way self-command will be achieved. Without disciplining his inclinations, man can attain to nothing... (*LE* 27:361).

Kant argues here that self-mastery requires weakening the forces opposed to it by “dividing” them. As I understand this process, dividing these forces is done in two stages: we first discipline ourselves and our inclinations, then moral feeling is cultivated. We discipline ourselves by rooting out the propensity that arises from the sensuous incentive. Kant says that rooting out this propensity is done by “repeated actions”, which suggests that self-discipline requires repeatedly acting as morality demands in the face of opposing inclinations. Moreover, self-discipline involves paying attention to ourselves, so that we often put our actions to the test

of "the inner judge" (conscience in its guise as practical reason, *MM*, 6:400-1; also called the "inner tribunal", *LE* 27:354) and examine their moral standing. So self-discipline requires moral self-scrutiny of one's actions and repeatedly performing one's duty in the face of opposing inclinations. And "performing one's duty" here should probably be understood in a broad sense, that includes not only performing morally required actions, but also abstaining from morally forbidden ones.

The second stage in which we "divide" the forces opposed to self-mastery is the cultivation of moral feeling. In opposition to self-discipline, which is a voluntary and reflective effort, cultivation seems to be a natural and involuntary consequence of self-discipline: we discipline ourselves and our inclinations by morally scrutinizing our actions and performing our duty. Then, "by long practice", we "will have given strength to the moral motive" and acquired a habit of experiencing pleasure in what is morally good and displeasure in what is morally evil,¹¹⁷ which Kant appears to be equating with moral feeling.¹¹⁸ As I understand this process, if we reflectively discipline our inclinations long enough, moral feeling will in time be involuntarily cultivated, so that we are habituated into taking pleasure in representations of morally required actions and displeasure in representations of morally forbidden ones.

To recap, discipline in Kant's generic sense is constraint by rules. Moral discipline consists in rationally constraining inclinations and the propensity to stray from moral laws on their account. It is done by scrutinizing our actions and repeatedly performing our duty

¹¹⁷ The Cambridge translation of *Lust oder Unlust* as "desire or aversion" is misleading, as it suggests Kant is talking about products of the faculty of desire.

¹¹⁸ Moral feeling as described in *MM* seems to be narrower than pleasure and displeasure in the morally good and evil; it is the susceptibility to feel pleasure or displeasure "merely from being aware that our actions are consistent with or contrary to the law of duty." (*MM* 6:399).

(including abstaining from forbidden actions), often in the face of opposing inclinations.

Cultivation in Kant's generic sense consists in acquiring or improving a faculty, skill or power for attaining rational ends - be they moral or non-moral. Cultivation of moral feeling consists in acquiring a habit of taking pleasure in morally required actions and displeasure in morally forbidden ones, and appears to be a natural consequence of disciplining our inclinations. More will have to be said in the following sections about how we cultivate and discipline feelings and inclinations. But if, as I have argued, Kant has this relatively stable discipline-cultivation distinction, a recommendation to discipline some of the mental states we call "emotions" and cultivate others is probably not accidental. To the contrary, I will argue in what follows that it is grounded in Kant's understanding of this distinction.

4.3. Cultivate Feelings, Discipline Inclinations

The primary textual evidence that Kant thinks feelings require cultivation while inclinations require discipline comes from a passage in the anthropology lectures of 1775-6, mentioned earlier:

- (a) The sensations of human beings require cultivation and the inclinations require discipline. The sensations should be refined and the inclinations brought under control. He whose sensations have not received cultivation, is unrefined, and he whose sensations do not admit of cultivation, is crude. (*LA* 25:621-2)

Kant is contrasting here sensations (*Empfindungen*) with inclinations: sensations require cultivation (*Cultur*) and should be refined, while inclinations require discipline (*Disciplin*) and should be brought under control (*gebändigt werden*). This contrast suggests that discipline is *not*

appropriate for sensations, and cultivation is *not* appropriate for inclinations. Moreover, since he is contrasting “cultivation” with “discipline”, I shall take Kant to be using “cultivation” here in the narrow sense of acquiring or improving a faculty, power, or skill for attaining rational ends (and not in the broad sense that includes discipline). Now, one might worry that this passage, taken from early lecture notes, does not capture what Kant actually said or does not reflect his mature view. I will return to these worries momentarily. First, we need to get clear on the nature of the objects that require cultivation and discipline according to this passage.

The object of discipline, inclination, in the usual sense this term has in Kant, is a habitual sensuous desire (*Anth.* 7:251; *MM* 6:212; *LA* 25:1334; and *Rel.* 6:29). In the previous chapter, we have seen that Kant at one point allows that we can speak of “sense-free” inclinations (*MM* 6:213). Here, however, the fact that Kant says we must discipline our inclinations and bring them under control strongly suggests that he is referring to sensuous inclinations, i.e., those which are caused by feelings of pleasure (elsewhere said to depend on sensations, *GMM* 4:413 fn.; and on feeling, *CPrR* 5:72-3).

The task of determining the object of cultivation is more complicated than that of determining the object of discipline because as noted in 2.2, Kant uses the term “sensation” to designate two different things: (1) A determination of the feeling of pleasure and displeasure, related solely to the subject, that does not provide cognitive material (*CPJ* 5:189, 206). “Sensation” in this sense is roughly-synonymous with “feeling”: it is the generic term which “pleasure” and “displeasure” (or “satisfaction and “dissatisfaction”) are “positive” and “negative” species of.¹¹⁹ (2) A representation introduced through sense that refers to a

¹¹⁹ Accordingly, Sorensen uses “feeling” and “sensation” as “more or less synonymous” (Sorensen 2002, 127, fn. 10).

modification in the subject, but can nonetheless provide material for cognition through its matter (*MM* 6:212, fn.; *Anth.* 7:153).

There are several reasons to think that in (a), “sensations” has sense (1). First, the examples Kant provides immediately after (a) are of refined feelings and sensitivities – not of sensations of the cognitive kind. The examples are a feeling of the duty of gratitude, being sensible to the indecency of harboring feelings of resentment, and esteeming a past friendship even after it is broken. Although they may depend on cognition of objects, these sensations and sensitivities do not contribute to cognition. Second, a bit earlier in the lecture, Kant says that keeping the mind’s composure with respect to feelings and desires requires preventing sensation from becoming affect, and desire from becoming passion. Affect is said there to be “That degree of sensation that makes us unable to estimate and compare the object with the sum total of all our sensation” (*LA* 25:590; cf. *Anth.* 7:254). Kant is therefore concerned with sensation as an influence on mental composure – not as a source of cognitive content. Finally, earlier in the lecture, Kant says that affect is a kind of feeling (“Affect is a feeling through which we lose our composure”, *LA* 25:589), and so the fact that he also says it is a degree of sensation strongly suggests that he is using “sensations” in (a) in the sense of “feelings”. I therefore submit that in (a), Kant is saying that the feelings of human beings require cultivation.¹²⁰

In the paragraph that follows (a), Kant speaks disapprovingly about “too great refinement of the sensations”. He argues that overly refined sensations in response to others’ needs in order to honor them can backfire and become a kind of flattery (*LA* 25:622). And he also praises an

¹²⁰ Earlier in the lecture, Kant speaks about cultivation of the senses, as when a hunter cultivates his vision (*LA* 25:500). He therefore allows that we can cultivate “cognitive” sensations. But for the abovementioned reasons, I believe that in (a) he is talking about feelings.

emotional sensitivity (*Empfindsamkeit*) that involves good judgment, and condemns an emotional susceptibility (*Empfindlichkeit*) of a different sort, which arises from overly refined sensations (cf. *Anth.* 7:236, where the contrast is between *Empfindsamkeit* and *Empfindlei*). But although he does not say in this paragraph that we ought to cultivate feelings or sensations, neither does he say we ought not do that. So the paragraph that follows (a) is compatible with the recommendation in (a) to cultivate sensations.

One might object that the fact that in the following paragraph he speaks against overly refined sensations and gives praise to emotional sensitivity suggests that what Kant meant to say in (a) is that we should cultivate emotional sensitivity, rather than sensations or feelings. But Kant clearly distinguishes between the sensations that are overly cultivated (sensations in response to others' needs) and the result of such exaggerated cultivation, i.e., emotional susceptibility. So there is no reason to think that the saying in (a) that sensations require cultivation does not represent Kant's view. We had better take Kant at his word: cultivate and refine sensations, but don't exaggerate, so that you do not become emotionally susceptible.¹²¹

In (a), Kant is therefore saying that we are to cultivate feelings and discipline inclinations. But does this accurately capture what he said in the lecture? The fact that he is recorded as expressing the same idea a bit later is evidence that it does:

(b) Just as cultivation aims at sensation, so discipline aims at the inclinations. He whose inclinations have not received any discipline is ill-mannered, however he

¹²¹ As we have seen in Chapter 2, Kant uses "feeling" to refer both to occurrent feelings of pleasure and displeasure and to the capacity for having them. If "sensation" is also ambiguous in this way, he may be saying in (a) that we are to cultivate the *capacity* for sensations (rather than occurrent sensations). A capacity for sensations, however, is not the same as emotional sensitivity: the latter involves judgment, while the former need not.

whose inclinations do not admit of any discipline is wild. In regard to the inclinations, the human being must be brought under control (*gebändigt*), just as in regard to the sensations, he must be refined. (LA 25:623)

Kant uses slightly different language here, saying that cultivation *aims* at sensation and discipline *aims* at the inclinations (rather than saying as he does in (a) that sensations *require* cultivation and inclinations discipline). But he is explicit that the human being's inclinations must be brought under control and that his sensations must be refined – the same treatments assigned to them in (a). And since cultivation aims at feelings and discipline at inclinations, it is reasonable to infer that Kant is making the same point as in (a), namely, that inclinations require discipline and feelings require cultivation (assuming “sensation” means “feeling”).

Moreover, later in this lecture, Kant argues that "Discipline is the constraint of inclination in accordance with rules" (LA 25:651), which suggests that *we cannot discipline anything but inclinations*; and he says that people whose inclinations have received no discipline are wild, but those "whose natural aptitude (*Naturell*) does not admit of any discipline at all ... are evil people." (ibid.) This suggests that inclinations (and natural aptitude) require discipline. And although he does not say later in the lecture that we are to cultivate feelings, he does say that "The individual who has received no cultivation is unrefined, the one incapable of any cultivation is crude." (ibid.). In light of (a) and (b), it is reasonable that cultivation here includes cultivation of feelings. If this is the case, Kant is saying that the individual whose feelings received no cultivation is unrefined and the one incapable of such cultivation is crude. So, (a) and (b) are probably accurate recordings of the view Kant presented in this lecture, dated 1775-6: feelings require cultivation, inclinations require discipline.

But did Kant stick to this prescription later in his career, in the 1780s and 1790s? There is evidence in the lecture notes that he did. Earlier we saw that in the *LE* of 1784-5, Kant says that "Without disciplining his inclinations, man can attain to nothing" (*LE* 27:361), and that he regards the discipline of inclinations as necessary for self-mastery. Although the first claim might be hyperbolic, the point seems to be that disciplining our inclinations is required not only for moral conduct but for prudential conduct as well. And since self-mastery is necessary for virtue (*MM* 6:383, 515), and disciplining inclinations is necessary for self-mastery, Kant must think we ought to discipline inclinations. In Busolt's notes of 1788-9, Kant says that "under 'discipline' one understands the limitation of inclination through a certain rule." (*LA* 25:1530). This is roughly the same characterization of discipline given some thirteen years earlier (*LA* 25:651) and further evidence that we cannot discipline other emotions. This claim appears in a discussion about acquiring character, which suggests that Kant thinks disciplining one's inclinations is a good and healthy process. Moreover, after presenting this conception of discipline, he says that "The human being assumes discipline when he himself denies it so far that he accommodates himself to others." (*LA* 25:1530). Although this falls short of a recommendation to discipline inclinations, it is reasonable to read Kant as saying that as we mature, we must assume discipline over ourselves such that we reject others' discipline. And since discipline has just been characterized as limitation of *inclination* through a rule, Kant is saying that as we mature, we must assume responsibility for disciplining our inclinations.

There is also evidence from a published work of the 1790s that we should discipline natural inclinations. Commenting on the stoic conception of virtue as presupposing the presence of an enemy, Kant says: "those valiant men [the Stoics] mistook their enemy, who is not to be sought in the natural inclinations, which merely lack discipline" (*Rel.* 6:57). Kant is not saying

here that we *ought* to discipline natural inclinations. But his point in saying that they “merely lack discipline” seems to be that although natural inclinations are not the enemy of virtue, virtue nonetheless requires their discipline. The next paragraph strengthens this impression:

Considered in themselves, natural inclinations are *good*, i.e., not reprehensible, and to want to extirpate them would not only be futile but harmful and blameworthy as well; we must rather only curb (*bezähmen*) them, so that they will not wear each other out but will instead be harmonized into a whole called happiness. (*Rel.* 6:58)

Kant says here that we must curb (*bezähmen*) natural inclinations, and there are several reasons to think he means we must discipline them: (1) In (a), he glosses “discipline” with “bringing under control” (*gebändigt*), and *bändig* is one of the meanings of “curb”(*bezähmen*);¹²² so, “curb” could mean “discipline”; (2) On the previous page (*Rel.* 6:57), Kant suggests that virtue involves disciplined natural inclinations, so it makes sense for him to proceed to say we must discipline them; (3) In *LP* Kant says that discipline is “the curbing [*Bezähmung*] of savagery” (9:449), and at *LA* 25:651, he seems to be using “curbing” and “disciplining” synonymously.¹²³ In light of this close connection between “curbing” and “disciplining”, it is reasonable to read a prescription to curb natural inclinations as saying we must discipline them. The discipline in question here, however, appears to be prudential (by rules of prudence) rather than moral, as Kant says that we should discipline natural inclinations so that “they will not wear each other out but will instead be harmonized into a whole called happiness.” And although the claim is about *natural* inclinations, it probably holds for all inclinations; for

¹²² See the Duden online dictionary (<https://www.duden.de/rechtschreibung/bezaehmen>).

¹²³ That “curb” and “discipline” there are roughly-synonymous is strongly suggested when Kant says “the human being is wild by nature; he has inclinations which simply take their course, if they are not tamed or curbed through art. Discipline is the constraint of inclination in accordance with rules.” (*LA* 25:561).

cultural inclinations, especially in their guise as passions for honor, possession and dominance, also pose a threat to morality and prudence (*Anth.* 7:267-274).

There is, therefore, strong evidence that we generally ought to discipline our inclinations. There is also evidence, however, that such discipline must not be exaggerated, and that we *can* cultivate inclinations, in the sense that it is within our power, and perhaps also *may* do so, in the sense that it is morally permissible. In *MM*, Kant says we have a duty to provide for ourselves "to the extent necessary just to find enjoyment [*Vergnügen*] in living"; and he warns that it is opposed to this duty to deprive oneself "of enjoyment of the pleasures of life by exaggerated *discipline* of one's natural inclinations." (*MM* 6:452). Kant's point here, I take it, is that we should not exaggerate in disciplining natural inclinations, lest we cease to find enjoyment in life. Moreover, as we have seen, Kant thinks that sensuous desires are caused by feelings of pleasure (*MM* 6:212), and so cultivating feelings probably has the effect of cultivating the inclinations caused by them. And the impression that it is possible to cultivate inclinations gains further support from passages where Kant speaks of refined inclinations (*CPrR* 5:39, 188; *CPJ* 5:297, 298). Since he sometimes uses "refined" synonymously with "cultivated" (*CPrR* 5:24, 39; *LA* 25:622), Kant's talk of refined inclinations suggests that inclinations can be cultivated, that is, that their cultivation is within our power.

Finally, there is one passage I am aware of where Kant appears to be saying that we may cultivate inclinations. In a discussion of duties of benevolence, he says that moralists "must lay down principles, and commend and cultivate the benevolent life based on obligation; and once all obligation has been set forth both through nature and religion [*alle Verbindlichkeit durch die Natur auch durch die Religion vorgelegt ist*], then the inclination, too, may be cultivated, but only insofar as it has to be subject to principles, and then they can be presented as incentives

[*Triebfedern*] to kindly actions from inclination." (*LE* 27:414-5). Kant's point here seems to be that the "moralists" - presumably, the moral philosophers or the teachers of morality - must first cultivate a benevolent life based on moral principles. Once this has been done (and once obligation has been properly presented), they may cultivate benevolent inclinations, but only insofar as they are subject to principles, which I understand as "disciplined". So, the cultivation of the benevolent life based on obligation consists in disciplining benevolent inclinations (not in their cultivation). But once they are disciplined, these inclinations may be cultivated. Cultivation of certain disciplined inclinations is thus permissible, according to this passage. This, however, is compatible with the recommendation that inclinations must – or at least must first – be disciplined.

Regarding Kant's view in the 1780s and 1790s on cultivating feelings, the evidence is not as decisive. There is, however, evidence that we should: (1) not weaken some feelings, other than the "moral endowments"; (2) increase the capacity for some pleasures (i.e., feelings that are not "moral endowments"); and (3) cultivate some feelings (which arguably are not "moral endowments"). Starting with (1), in *MM*, Kant argues that a propensity to wanton destruction of inanimate nature is opposed to a duty to oneself because it weakens or uproots a feeling of love for inanimate things apart from an intention to use them (*MM* 6:443). This means that we should not weaken this feeling of love, which is not one of the "moral endowments".

In *Anth.* there is evidence that we should increase the *capacity* for feelings that are not "moral endowments":

One way of enjoying ourselves is also a way of *cultivating* ourselves; that is, increasing the capacity (*Fähigkeit*) for having more enjoyment of this kind, and this applies to the sciences and the fine arts. However, *another* way is *overindulgence*, which makes us increasingly less capable of further enjoyment. (*Anth.*7:236-7)

Kant is saying here that enjoyment could be a way of cultivating ourselves, when we increase our capacity to have more enjoyments of a certain kind. Enjoyment (*Vergnügen*), however, is said a bit earlier to be pleasure introduced through sense (*Anth.* 7: 230), and later is explained as a conscious feeling that urges the subject to remain in his state (7:254). So Kant's point seems to be that by enjoying ourselves, that is, by having sensuous pleasures, we can cultivate ourselves and increase our capacity for having more of those pleasures. And since the capacity for pleasure is called "feeling" (*MM* 6:211), and enjoyment is a feeling, Kant is saying that we can cultivate certain feelings. To use his examples, by enjoying scientific inquiry, I can increase my capacity for pleasure in science; and by enjoying a Rembrandt portrait, I can increase my capacity for pleasure in fine arts.¹²⁴

Right after the abovementioned passage, Kant warns about the dangers of overindulgence: "- - Young man! ... get fond of work; deny yourself enjoyments, not to *renounce* them, but rather to keep them always in perspective as far as possible!" (*Anth.* 7:237; cf. 165). This warning against overindulgence and the injunction to deny oneself enjoyments might seem like a recommendation to *discipline* sensuous pleasures. But note that Kant prescribes *growing fond* of work. Given that he has just said that enjoying oneself is a way of increasing one's capacity for enjoyment, it is reasonable to read him as prescribing increasing the capacity to enjoy work, i.e., to take pleasure in it. We do so, according to this passage, by denying ourselves sensuous pleasures when we are young, because overindulging in them might lead to desensitization and decreasing our capacity for having them in the future. And given that

¹²⁴ Kant's examples do not seem to involve sensuous pleasures, but rather an intellectual pleasure (in the sciences) and a partly-intellectual partly-sensuous pleasure (in fine arts). This suggests that we can cultivate feelings of all three kinds: sensuous, intellectual, and partly-intellectual partly-sensuous.

in *LE* Kant is recorded as saying that "the greatest discipline is to accustom oneself to work" (*LE* 27:396), it is reasonable that increasing enjoyment in work also involves disciplining "lazy" inclinations. But the fact that Kant prescribes that young people deny themselves enjoyments need not mean that they should *discipline* their feelings; it rather suggests that increasing capacities for pleasures may involve denying ourselves pleasures.

There is also evidence from *MM* that certain feelings should be cultivated. First, Kant argues that gratitude is a duty and consists in "*honoring* a person because of a benefit he has rendered us. The feeling connected with this judgment is respect for the benefactor" (*MM* 6:454). He goes on to say that "Even mere heartfelt *benevolence* on another's part, without physical results, deserves to be called a duty of virtue; and this is the basis for the distinction between *active* and merely *affective* gratitude." (*MM* 6:455). The analogy between heartfelt benevolence and affective gratitude suggests that the latter too, in its guise as mere feeling without "physical" results, is a duty of virtue. And if this is the case, it is reasonable that we should cultivate this feeling (as is also suggested by Kant's citing the feeling of the duty of gratitude after his recommendation to cultivate sensations in (a)). It might be thought that because the feeling of gratitude is connected with respect, it is one of the "moral endowments". But respect as a moral endowment is *self-respect* (*MM* 6:402-3), not respect to another person, and not specifically to a benefactor.

Compassionate feelings also require cultivation, according to the following famous passage:

But while it is not in itself a duty to share the sufferings (as well as the joys) of others, it is a duty to sympathize actively in their fate; and to this end it is therefore an indirect duty to cultivate the compassionate (*mitleidige*) natural (aesthetic) feelings in us, and to make use of them as so many means to sympathy based on

moral principles and the feeling appropriate to them. It is therefore a duty not to avoid the places where the poor who lack the most basic necessities are to be found but rather to seek them out, and not to shun sickrooms or debtors' prisons and so forth in order to avoid sharing painful feelings one may not be able to resist. For this is still one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish. (*MM* 6:457)

Kant is saying here that we have an indirect duty to cultivate compassionate natural (aesthetic) feelings. What the duty to cultivate compassionate feelings involves is a matter of dispute. Thomason, for instance, argues that in saying that we should cultivate sympathetic feelings, "Kant might just be saying: do not try to get rid of them or see them as a burden." (Thomason 2017, 446). But even she grants that at 6:457, Kant "explicitly states that it is our "indirect duty" to cultivate our natural feelings of sympathy." (ibid., 444).¹²⁵ So Thomason concedes that in this passage, Kant recommends cultivating feelings. Moreover, in saying we should cultivate compassionate feelings, Kant cannot *just* be saying that we should not try to get rid of these feelings and not see them as a burden. For he is explicit that we have a duty to (a) not avoid the poor who lack the most basic necessities; (b) seek out these poor people; and (c) not shun sickrooms or debtors' prisons. Thomason argues that we should seek out places where people who need help are found "because the people there need our help, but not because they are good places to practice sympathy." (ibid., 446). But Kant's use of "therefore" after the sentence about the duty to cultivate compassionate feelings makes clear that it is *because* we

¹²⁵ Thomason's claim is not entirely accurate because Kant says we have an indirect duty to cultivate *compassionate* (*mitleidige*) natural feelings, not sympathetic (*theilnehmende*) feelings (which Fahmy proposes to translate as "participatory", to emphasize the active nature of the duty regarding them, 2009, 36). But Kant explicates "shared feeling" (*Mitgefühl*) as "sympathetic sensation" (*theilnehmende Empfindung*), and so "compassionate" and "sympathetic" seem to be roughly equivalent. Thomason too, I believe, is using "compassionate" and "sympathetic" as roughly equivalent.

have this duty that we are obligated to (a)-(c). So, although the reason we have a duty to cultivate compassionate feelings is that we should "sympathize actively in their fate", the reason we ought to seek out the poor is that we ought to cultivate compassionate feelings. Granted, the cultivation of these feelings is a means to active sympathy; but their cultivation requires more than not getting rid of them and not seeing them as a burden.

The above passage therefore makes clear that we have a duty to cultivate compassionate feelings, and that this requires seeking out the poor, not avoiding them, and not shunning sickrooms and debtors' prisons. It is less clear, however, what Kant means when he says we have a duty to use those feelings as "means to sympathy based on moral principles and the feeling appropriate to them". One way to understand this duty is that we ought to use compassionate feelings to *motivate* acts or expressions of sympathy. This is suggested by Kant's saying that painful feelings in response to others' suffering are "one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish." Kant might therefore be recommending that we use compassionate feelings to perform sympathetic actions in conformity with duty, but not done *from* duty, when the motive of duty does not move us to perform these actions.¹²⁶ Considerations regarding Kant's theory of moral motivation might lead us to look for an alternative reading, on which compassionate feelings do not motivate the actions they are means for.¹²⁷ The non-motivational function of compassionate feelings could be

¹²⁶ This appears to be Borges' view, as she argues that "The cultivation of sympathy seems to fulfill the role of a moral incentive when the law is not sufficient to promote the moral action." (Borges 2019, 79).

¹²⁷ If Kant's view were that (1) there is something morally amiss in performing our duty not *from* duty, and (2) whether an action is done from duty is determined by its particular incentive, rather than by the incentive incorporated into the agent's fundamental maxim of character, he probably would not recommend cultivating compassionate feelings as incentives to morally-required action (although he might nonetheless recommend this practice, given that performing one's duty not from duty is better than not performing one's duty at all). If he is not committed to both (1) and (2), I see no reason to avoid attributing this recommendation to Kant.

perceptual sensitivity to situations where we can help others (Baron 1995, 171), facilitating communication of these feelings and communication informed by them (Fahmy 2009, 42), or expression of morally significant attitudes toward others (Baxley 2010, 165; Sherman 1997, 147). Kant's saying that we are to use them as means for "sympathy based on moral principles and the feeling appropriate to them" suggests that the acts which sympathetic feelings are means for are motivated not by these feelings but rather by moral feeling (at the thought of one's duty, rather than of the plight of others). But independently of whether compassionate feelings are supposed to motivate acts of sympathy, *MM* 6:457 makes clear that we should cultivate certain feelings as means for sympathetic actions.

But are compassionate feelings different from "moral endowments"? They do seem very similar to "love of human beings" (*Menschenliebe*, earlier referred to as "love of one's neighbor", *MM* 6:399); for love of human beings is characterized as "an aptitude (*Fertigkeit*) of the inclination to beneficence in general" (*MM* 6:402), and compassionate feelings too facilitate acts of beneficence. There are two reasons, however, to think that love of human beings in its guise as a moral endowment is different from compassionate feelings. First, because as a moral endowment, love of human beings is a condition of our receptivity to concepts of duty as such, such that it is by virtue of having it that we can be put under moral obligation (*MM* 6:399). In other words, not having one of the four moral endowments would mean that our moral agency is impaired in a way that exempts us from certain duties, and perhaps even from *all* duties. In contrast, lacking in compassionate feelings, although it would make the fulfillment of acts of beneficence difficult, does not seem to exempt one from the duty of beneficence (and certainly not from all duties). So, compassionate feelings differ from love of human beings in that our

obligation to certain moral duties (and perhaps to all of them) depends on our ability to experience the latter, but not on our ability to experience the former.

The second reason to think love of human beings is not the same as “compassionate feelings” is that Kant's Latin term for love of human beings, as an immediate pleasure in an object's existence, is *amor complacentiae* (*MM* 6:402); the Latin term he equates with compassion, in contrast, is *communio sentiendi illiberalis, servilis* (*MM* 6:456-7) (“unfree communion of sentiment, slavish”).¹²⁸ So, compassionate feelings are probably different from “love of human beings”. I therefore submit that in the 1790s, Kant recommends increasing capacities for pleasures that are not moral endowments, and that he prescribes cultivating feelings that are probably not moral endowments.

But although on my reading Kant prescribes that we cultivate feelings, his considered view is probably that although they are feelings, affects should not be cultivated. Affects, as we have seen in 1.4, are feelings that do not admit of reflection (*Anth.* 7:251). They pose a serious threat to rational deliberation, making rational reflection difficult or impossible (*ibid*; cf. *MM* 6:407). And although Kant admits that affects can be used for promoting rational ends, he also warns that considered by itself alone, affect “makes itself incapable of pursuing its own end, and

¹²⁸ Schönecker agrees that love of human beings as a moral endowment is different from compassionate feelings (Schönecker 2010, 157). In contrast, Baron's view appears to be that the susceptibility for compassionate feelings is either identical with or narrower than the moral endowment of love of human beings (or love of one's neighbor); for she argues that *humanitas aesthetica*, i.e., the natural susceptibility to feel joy and sadness in common with others, is either identical with or narrower than “love of one's neighbor” (Baron 1995, 212). Although she does not say why this is the case, it is reasonable to understand compassionate feelings as narrower than or identical with love of human beings if we understand the latter as pleasure in other human beings' existence (as suggested by Kant's saying that the love that is satisfaction is “a pleasure joined immediately to the representation of an object's existence” *MM* 6:402), and the former as pleasure and displeasure in sync with another's. But elsewhere, Kant's view appears to be that love of human beings as a feeling is pleasure in their *perfection* (*MM* 6:449). And the ability to feel pleasure and displeasure in sync with others' is neither identical nor narrower than the ability to take pleasure in their perfection.

it is therefore unwise to allow it to come into being intentionally." (*Anth.* 7:253). Cultivating affects, since it may involve seeking occasions to have them, may involve intentionally allowing them to come into being; and so their cultivation would be "unwise" and involve actively undermining rational deliberation. Kant would therefore not recommend cultivating affects.¹²⁹

Should affects be disciplined, then? Some passages suggest this much. In the *Anth.*, for example, Kant cites as a rule "for a tasteful feast" that in a serious conflict that cannot be avoided, one must "maintain discipline over oneself and one's affects [*Affect*], so that mutual respect and benevolence always shine forth" (*Anth.* 7:281); and in *MM*, he says that inner freedom, which virtue is based on (*MM* 6:408), requires *subduing* one's affects (*seine Affecten zu zähmen*, *MM* 6:407). But notice that in the first of these two passages, Kant says one should maintain discipline *over* one's affects, which is not quite the same as disciplining the affects themselves. And in the second passage, after he says that inner freedom requires subduing one's affects, he adds that it also requires that one's passions be *governed* (*seine Leidenschaften zu beherrschen*, *MM* 6:407). Since passion is a kind of inclination, it is reasonable that here too, the treatment assigned to affects (which are feelings) differs from the one Kant would assign to inclinations. I therefore propose that when he says that we are to maintain discipline *over* affects and that inner freedom requires subduing our affects, Kant means that we should discipline the inclinations caused by them. And so, although in (a) and (b) he recommends cultivating feelings,

¹²⁹ This does not mean that affects are bad in every respect. As Baron notes, Kant says that some affects can be conducive to health (Baron 1995, 200). Moreover, Kant's attitude toward affects (and toward feelings in general) seems to depend on the origin of the representations that give rise to them (whether these are intellectual representations of reason and the understanding, or sensuous representations of sensibility); and he certainly sees the value of affects caused by representations of reason, such as enthusiasm, astonishment and fortitude (see Sorensen 2002, Clewis 2009). But for the above reasons, I do not think he would recommend cultivating affects.

Kant's considered view is probably that we should cultivate feelings except for affects; and that affects should be subdued by disciplining the inclinations they give rise to.

4.4. Why Different Treatments?

If what I have argued is correct, Kant's view is that feelings except affects require cultivation (i.e., acquiring or improving them so that they can be used for promoting rational ends), and inclinations generally require discipline (i.e., constraint by means of rules or laws). One worry regarding this reading is that Kant should not recommend cultivation of any feelings or desires, or the cultivation of any feelings or desires other than those that stem from our "moral endowments". I started dealing with this worry in the previous section, and I will revisit it in the next. Before I do that, however, there are two other worries which need to be addressed. The first is that a recommendation to cultivate feelings and discipline inclinations is implausible because it entails that we are to cultivate feelings that are opposed to moral conduct, and discipline inclinations that are not opposed to it or even facilitate it.

In response to this worry, it must be conceded that Kant does seem to think that certain feelings and desires are usually conducive to moral conduct (for example, those involved in the duties of beneficence, gratitude and sympathy, *MM* 6:452-8), while others (e.g., those involved in the vices of envy, ingratitude and malice, *MM* 6:458-62) are usually opposed to it. But it is doubtful that any feelings or desires - with the exception of those that stem from our "moral endowments" and of passions¹³⁰ - are *necessarily* or *in principle* conducive or detrimental to

¹³⁰ The caveat regarding passions is meant to accommodate Kant's saying that passions are "without exception evil" (*Anth.* 7:267), which I take to mean that passions *are* necessarily detrimental to morality. The reason that passions are necessarily detrimental to morality is that they involve an evil maxim. As Kant explains in *MM*, while affects precede reflection, passions involve a calmness that "permits

morality. For although taking pleasure in certain kinds of objects causes inclinations that are usually conducive to morality (and taking displeasure in other others causes disinclinations usually detrimental to it), it is doubtful that there is an object pleasure in which would cause an inclination that is always conducive or detrimental to morality. For example, sympathetic sadness, i.e., displeasure in another's pain (*MM* 6:456), is generally conducive to morality in that it causes an inclination to help others promote their happiness (which is a duty of wide obligation, *MM* 6:393).¹³¹ But - to borrow Herman's example (Herman 1993, 4-5), my sympathetic sadness at the plight of a person struggling with a heavy burden at the backdoor of the Museum of Fine Arts is detrimental to morality insofar as it inclines me to help a thief. Similarly, my envious feeling at a colleague's success and the resultant disinclination to help her are detrimental to morality insofar as they dispose me not to help her promote her happiness. But they are also conducive to morality insofar as they dispose me to cultivate my talents and promote my natural perfection (which is a duty of wide obligation, *MM* 6:391-2).¹³²

If this is correct, then Kant's unqualified recommendation that we cultivate feelings and discipline inclinations suggests that we should cultivate even those feelings that are usually opposed to morality and discipline inclinations that are usually conducive to it. This, however,

reflection and allows the mind to form principles upon it and so, if inclination lights upon something contrary to the law, to brood upon it, to get it rooted deeply, and so to take up what is evil (as something premeditated) into its maxim." (*MM* 6:408). That passions involve maxims might lead us to think that they are not emotions, if we think that emotions cannot be "principled" in this manner. But hatred is a passion, according to Kant (*MM* 6:408), and so, if hatred is an emotion, at least some passions, despite being "principled", are emotions.

¹³¹ More accurately, we should say that sadness at others' pain causes a *disinclination* not to help them. For Kant's view appears to be that desires (+) and inclinations (+) are causally connected to pleasures (+), while aversions and disinclinations are causally connected to displeasures (*MM* 6:211-2).

¹³² Sorensen makes a similar point regarding inclinations, saying that Kant's view appears to be that they are not reliable as a criterion for morality or as a motivation for it (Sorensen 2002, 111). I add that inclinations (and feelings) are not reliable in these ways because they are neither necessarily conducive nor detrimental to morality.

does not entail that we should cultivate feelings that are *always* opposed to morality, and neither does it entail disciplining inclinations that are always conducive to it, for there simply are no such feelings and inclinations (aside from those grounded in our moral endowments, and passions). And although it might be thought implausible that we ought to cultivate feelings that are usually opposed to morality, the recommendation makes sense if such feelings could be used to promote rational but not necessarily moral ends (as seems to be the case with envy, for instance); and a recommendation to discipline inclinations that are usually conducive to morality makes sense if discipline involves constraint by means of rules, which is necessary in case a usually-good inclination prompts us to perform a forbidden action or not to perform an obligatory one.

The other worry is that it is implausible that Kant would assign different treatments to inclinations and feelings because the difference between them is not significant enough to warrant different treatments. This worry might seem to gain support from the many qualifications I have introduced to Kant's original recommendation (that inclinations can be cultivated and should not be overly-disciplined, that affects should not be cultivated, etc.). And it also seems to be supported by the fact that inclinations depend on feelings and are caused by them, and that they both seem to be involuntary expressions of our sensuous nature. Given their close connection and similarity, there seems to be no reason that inclinations should be treated differently from feelings.

Recall, however, that earlier in this chapter we saw that discipline involves constraining a tendency to deviate from a certain rule or law. And if what I have just argued is correct, both feelings and inclinations may lead us to deviate from moral principles, as neither of them necessarily disposes us to perform our duty. But we had also seen in the previous chapter that

while inclinations (and desires in general) can motivate action immediately, feelings must cause desires or inclinations if they are to produce action. I therefore propose that inclinations (and desires in general) should be disciplined because discipline involves preventing a mental state from motivating an action prohibited by a rule or a law (or preventing it from motivating the omission of an action required according to a rule or a law), and inclinations can motivate such actions and omissions immediately; feelings, in contrast, cannot be disciplined because although they cause inclinations that may lead to deviating from laws or rules, they cannot motivate actions (or omissions) immediately. And since feelings nonetheless can be used to promote rational ends, they ought to be cultivated.

To see how this model of discipline works, think of a case where you are tempted to take money from a stranger's wallet. To the extent that the thought of this action tempts you, you feel pleasure at the thought of performing it or at the thought of its expected outcomes (say, getting that new pair of jeans). This pleasure, in turn, causes an inclination to perform the immoral action. Discipline, on my proposal, is supposed to occur at this stage in the process, after the inclination appears; it is then that you can think of the moral principle that prohibits stealing (or prohibits stealing in cases of this kind), and then choose not to act on your inclination. Such a choice would of course mean that your feeling of pleasure at the thought of the immoral action would fail to bring it about. But this choice involves constraint of inclinations, not of feelings, because it prevents your inclination from motivating action but does not prevent your feeling from causing the inclination.¹³³

¹³³ There is a sense, of course, in which both the feeling and the inclination “motivate” the action even when they fail to bring it about; for even in this case, we are moved toward taking a certain action. But an inclination or feeling that fails to bring about action does not motivate in the sense that the action toward which it moves us is never taken.

To this proposal one might object that neither our feelings nor our inclinations can be constrained, according to Kant, and so he cannot recommend that we discipline our inclinations, if this involves constraining them. That feelings cannot be constrained is suggested when Kant says that "*Love* is a matter of *sensation* [*Empfindung*], not of willing, and I cannot love because I *will* to, still less because I *ought* to (I cannot be constrained to love); so *a duty to love* is an absurdity." (*MM* 6:402). Kant is saying here is that is impossible to be constrained to love (*zur Liebe genöthigt werden*), i.e., to have a sensation of love; and although the verb he uses is derived from *Nöthigung* rather than *Zwang*, he uses these terms as synonyms ("every duty is *necessitation*, a constraint", *MM* 6:402; and in *MM* 6:379). Kant says here that we cannot be constrained to love because love is a matter (*eine Sache*) of sensation, and this suggests that *all* sensations are outside the reach of constraint. Assuming that Kant is using "sensation" here in its non-cognitive sense, his point is that we cannot be constrained to have or experience feelings. And he makes a similar point about love as an inclination ("love as an inclination cannot be commanded", *GMM* 4:399). So, his view appears to be that we cannot be constrained to have or experience feelings or inclinations.

In response, it is important to first note that it is actually not clear whether the point of these passages is that feelings and inclinations cannot be constrained or that feelings and inclinations *of love* cannot be constrained. The claim that "*Love* is a matter of *sensation* [*Empfindung*], not of willing" suggests that all feelings and inclinations, as "matters of sensations", cannot be constrained; but Kant also says that "What is done from constraint... is not done from love" (*MM* 6:401), which suggests that the reason that we cannot be constrained to

love is not that it is a sensation, but rather that it is *love*; and if this the case, it is possible that non-loving feelings and inclinations *can be* constrained.¹³⁴

But even if Kant's point holds for all feelings and inclinations, such that we cannot be constrained to *have or experience* feelings or inclinations of all kinds, this does not entail that we cannot constrain our inclinations in the sense of preventing them from motivating prohibited actions or preventing them from motivating the omission of required actions. And Kant is explicit that human choice is affected but not determined by inclinations or sensible impulses (*MM* 6:213), such that we can choose whether to act on our inclinations. Our feelings and inclinations (or our loving feelings and inclinations) are therefore similar insofar as we cannot be constrained to have either; they differ, however, in that we can constrain our inclinations (but not our feelings) such that they do not motivate prohibited actions or the omission of required ones.

4.5. Defending the Cultivation of Feelings

In 4.3, I adduced evidence that Kant thinks we should cultivate some feelings and increase the capacities for them. One might still worry, however, that some of the things Kant says preclude a

¹³⁴ It should be noted, however, that the claim that "What is done from constraint... is not done from love." (*MM* 6:401) need not be read as saying we cannot be constrained to feel love or have a loving inclination; it may instead be read as saying that we cannot be constrained to do what we are inclined to do (or "feel like" doing). And this does not apply to loving inclinations alone, for I cannot be constrained to hurt someone I hate, if so inclined. As Kant explains earlier, constraint consists in doing something reluctantly, i.e., in the face of opposition from inclinations (*MM* 6:379). If this is the case, a constraint to perform an action one is inclined to is contradictory: it amounts to performing an action in the face of opposition from inclinations when there is no such opposition (assuming that being inclined to perform an action rules out opposition from one's inclinations). But even when read in this way, Kant's point is compatible with a recommendation to discipline inclinations; for although we cannot constrain ourselves to perform actions we are inclined to, we can constrain our inclinations by preventing them from motivating prohibited actions (and by performing required actions when inclined otherwise).

recommendation to cultivate feelings. I now wish to deal with this worry by considering Thomason's recent objections to the view that Kant could recommend emotional cultivation.¹³⁵ Thomason distinguishes between three ways of understanding what cultivation (of emotions) means in Kant: "refining", "strengthening", or "reflecting upon". Of these, my understanding of cultivation is closest to "refining", and so I will focus on her objections to this reading.

According to Thomason, the possibility that cultivation means "improve" or "refine" is suggested by Kant's saying that the duty of self-perfection consists in part in cultivating one's faculties or natural predispositions, which involves raising ourselves from the crude state of nature toward humanity (*MM* 6:387). And she adds that on this reading, in section 35 of *MM*, Kant seems to suggest that our natural sympathetic feelings are "the raw material that we refine into active sympathy. Indeed, he claims that active sympathy is the "end" in the service of which we should "cultivate the compassionate natural (aesthetic) feelings in us" (*MM* 6:457)." (Thomason 2017, 446). The problem Thomason sees for this reading is that natural compassionate feelings, according to Kant, belong to our receptivity, may be irresistible, are natural, pathological and unfree, and communicable like a contagious disease. If this is the case,

¹³⁵ Thomason's understanding of emotion is based on what Kant says of *Gemüthsbewegung* (Thomason 2017, 451), and Kant's view appears to be that *Gemüthsbewegungen* are affects, or affects and passions (*MM* 6:407-8; 25:589, 1340). One might therefore think that when Thomason argues that Kant does not believe we should cultivate emotions, she means he does not believe we should cultivate affects and/or passions - not that we should not cultivate *any* feelings or desires. But Thomason's reasons to reject the cultivation of emotions do not rely on her account of emotions; and she makes clear that by "emotion" she does not mean affects alone when she argues that affects are a subset of emotions that are more disruptive of inner freedom (Thomason 2017, 452). I therefore think that Thomason's view is that Kant does not believe we should cultivate any of our feelings and desires, with the possible exception of compassionate feelings (but their cultivation, according to Thomason, is limited to not trying to get rid of them and not seeing them as a burden).

it is unclear how such feelings can be refined into active sympathy, which is a free capacity or faculty (*Vermögen*).

Thomason is correct insofar as Kant takes the cultivation of natural compassionate feelings to be a means to active, rational sympathy. But Kant neither says nor suggests that such cultivation turns compassionate feelings into a free rational capacity.¹³⁶ He rather says that these feelings can be used as "so many means to sympathy based on moral principles and the feeling appropriate to them." (*MM* 6: 457). Earlier I argued that it is unclear whether feelings used in this way play a motivational role in performing morally required actions. But either way, compassionate feelings need not be understood as *turning into* a free rational capacity or active sympathy. They are better thought of either as motivating morally required action (that is not performed from the motive of duty) or as helping us perform acts of active sympathy without motivating them (by making us sensitive to others' needs or express morally significant attitudes).

The other major objection Thomason raises concerns the supposed-incompatibility of a recommendation to cultivate emotions with Kant's warning of the dangers of moral enthusiasm. Thomason argues that Kant "specifically warns" that cultivating a kindly disposition of spontaneously and gladly doing what is morally required without the need to constrain ourselves "risks locating 'the incentive [of morality] *pathologically* (in sympathy or self-love)' (*CPrR* 5:85,

¹³⁶ Granted, in *LP* Kant says that cultivation involves acquiring a faculty (*Vermögen*) "sufficient for the carrying out of whatever purpose." (*LP* 9:449). So, cultivation of feelings involves acquiring a faculty, and on my reading, this faculty is used to promote rational ends. But the process of cultivation need not be understood as one where the thing being cultivated turns into a faculty. It is rather better understood as a refinement of a receptivity or an aptitude (e.g., compassionate feelings), which leads to acquiring a faculty (performing beneficent actions).

emphasis in original). We are tempted to imagine ourselves as so good-hearted that we do not need the 'yoke' of duty (ibid.)." (Thomason 2017, 448). Thomason adds that if Kant thinks that the danger of believing we do not need to constrain ourselves to act morally is real, it would be odd of him to recommend that we intentionally expose ourselves to it. Anticipating a possible objection, she says that it may be thought there is no such danger because we can cultivate natural compassionate feelings and still act from duty and not from these feelings. And she responds that, although this may be possible, it seems unlikely that Kant would recommend engaging in behavior that will blur the boundaries between the motive of duty and that of sympathy.

To assess this objection, let us have a look at the passage Thomason thinks provides a warning against cultivating a kindly disposition:

By *exhortation* (*Aufmunterung*) to actions as noble, sublime, and magnanimous, minds are attuned to nothing but moral enthusiasm and exaggerated self-conceit... For, *when they imitate* such deeds - namely, from such a principle... they produce in this way a frivolous, high-flown, fantastic cast of mind, flattering themselves with a spontaneous goodness of heart that needs neither spur nor bridle and for which not even a command is necessary and thereby forgetting their obligation, which they ought to think of rather than merit. (*CPrR* 5:84-5, italics mine)

Kant's warning here is not of the dangers of cultivating our feelings or inclinations. Rather, it is of the dangers of a certain way of teaching morality that aims to promote moral conduct by encouraging the performance and imitation of actions presented as noble and sublime. The actions presented in this way and imitated are done on a principle of *merit*. As Kant explains in *MM*, an action has merit when the agent does more than she could be constrained to by law (*MM* 6:227). There is, of course, no harm in acting on such a principle when an action is not morally obligatory and is in fulfillment of an imperfect duty. But Kant seems to think that if

we emphasize the noble nature of these actions rather than focusing on morally required or prohibited actions, or if we try to imitate such actions, we will end up deceiving ourselves that we are beyond the need for self-constraint, such that nothing we do is morally required of us – it is all done out of the goodness of our heart. And this will, in turn, lead us to think that we no longer need to exercise self-constraint in order to act morally.¹³⁷

When I cultivate my feelings, however, I am not necessarily imitating actions done on a principle of merit, and I am also not necessarily encouraging anyone to perform such actions. Rather, such cultivation may involve not avoiding and seeking out opportunities to have feelings so that they can be used to pursue rational ends and perform my moral duty. Granted, the cultivation of certain feelings, such as sympathy, may make us more vulnerable to moral enthusiasm. For if our feelings are naturally in tune with those of others, such that we take pleasure in their fortune and displeasure in their misfortune, we might be tempted to think that we are beyond the need for self-constraint in performing our duties. But in the above passage, Kant is not talking about this danger, and there is no reason to think it is serious enough to warrant a prohibition on cultivating feelings. The fact that he recommends cultivating certain feelings suggests that the benefits of such cultivation for moral conduct outweigh its potential harms.¹³⁸ I therefore submit that despite attributing to Kant the view that we should cultivate feelings, my reading is not vulnerable to Thomason's pertinent objections.

¹³⁷ Commenting on this passage, Baron adds that in addition to the danger that people would not see themselves as constrained by morality, there is also the danger that they will "see themselves as constrained, but only because of their merit – a version of noblesse oblige." (Baron 1995, 37).

¹³⁸ It might be objected that if cultivation of feelings may involve seeking opportunities to have them, it may involve strengthening them. If this is correct, my account is also vulnerable to Thomason's objection to the reading of "cultivation" as "strengthening", namely, that cultivation would increase our passivity and decrease our freedom, which is the opposite of what morality requires (Thomason 2017, 447-8). A possible response to this worry is that Kant does not think that *all* feelings increase our passivity and

To recap, if what I have argued is correct, there is no correct “yes or no” answer to the question “Should Kantians cultivate their emotions?”. Assuming, as we have, that what we call “emotion” refers in Kant to mental states that originate in the faculty of desire and in feeling, the answer depends on the kind of feeling or desire we are talking about, on its impact on rational deliberation, and on its ability to motivate action immediately. Inclinations should generally be disciplined, i.e., constrained so that they do not motivate the performance of prohibited actions or omission of required ones, because they immediately motivate actions that are not necessarily the ones required. Feelings, expect affects, ought to be cultivated because they do not motivate actions immediately, can be used for pursuing rational ends, and do not significantly undermine rational deliberation.

decrease our inner freedom - only affects do. So, a recommendation to cultivate feelings (other than affects) is not opposed to morality. The rejoinder might be, however, that if cultivation strengthens our feelings, it is not clear we would be able to stop before they turn into affects. I therefore think that a better response is that although cultivation of feelings may involve seeking opportunities to have them, this need not lead to strengthening our feelings such that they become more intense; it may rather lead to our becoming more susceptible to having them or not being overwhelmed by them. To this one might object that if Kant's prescribed treatment for compassionate feelings, for example, involves this much, this treatment involves desensitization, such that the suffering of others does not bother us as much as it used to; and such a treatment cannot be said to be cultivation (this is the point Thomason appears to be making at 2017, 447). The objection is correct insofar as cultivation of feelings cannot lead to *indifference*, but I do not see evidence that it cannot lead to their becoming less intense and less overwhelming; and I do not think that the ordinary use of the word precludes this much. My loving feelings for an old friend, for example, may be cultivated not in the sense that they have become stronger over the years but rather in the sense they have become more deeply entrenched and stable.

Conclusion. Two Theories of Emotion?

On the reconstruction the foundations for which I have provided here, Kantian emotions come in two kinds: feelings, and desires. Emotions that are feelings are representations of objects' relations to a subject (as agreeing or disagreeing with its ability or intention to act), that have a felt phenomenological quality, and dispose their subject to preserve its state or change it, or to act in order to produce or not to produce an object. Emotions that are desires are dispositions or motivations to action that are causally connected with practical feelings, but differ from them in that desires (except for certain wishes) are directed at future objects, while feelings are directed at the subject's affective state. It is the future-directedness of non-wishful desires that allows them to motivate action immediately. And it is because emotions that are desires (and inclinations in particular) can motivate action immediately, that Kant thinks we should discipline inclinations, i.e., constrain them by rules so that they do not motivate morally or prudentially prohibited actions, and so that they do not motivate omissions of morally or prudentially required actions. In contrast, emotions that are feelings - except for affects - should be cultivated because they cannot be constrained by rules or laws (in the sense of preventing them from immediately motivating prohibited actions or the omission of required ones), but can be used in the pursuit of rational ends, and do not significantly undermine rational deliberation.

This does not necessarily amount to a complete reconstruction of Kant's theory of emotion because although I take myself to have provided the necessary conditions on a mental state's being a Kantian emotion, these conditions may be insufficient. If some but not all Kantian desires are emotions, then my account thus far provides the necessary but insufficient conditions on a mental state's being an emotion of the desire kind; and similarly, if some but not all Kantian feelings are emotions, it provides the necessary but insufficient conditions on a mental state's

being an emotion of the feeling kind. Providing the necessary *and* sufficient conditions on a mental state's being a Kantian emotion, however, would require introducing further extra-textual assumptions, which Kant might reject, about which mental states designate emotions or about the nature of emotions. There is of course no harm in adding such assumptions in order to try and get a more complete reconstruction; but I think it is best to do that *after* we have a clear picture of the nature of the genera to which Kantian emotions belong, so that we can clearly distinguish between what Kant says, and what he would or should have said, had he subscribed to our understanding of emotions.

For example, if we added to my account the extra-textual assumption that "pain" and "laughter", and perhaps other terms for sensuous feelings through sense (such as "pleasure in the taste of chocolate", or "pleasure in a soft kiss") designate mental states too primitive to qualify as emotions, then not all Kantian feelings qualify as emotions; for as we have seen in Chapter 2, laughter and pain are sensuous pleasure (in its guise as affect) and displeasure through sense, respectively, and all pleasures and displeasures of this sort are feelings. Similarly, if we added the assumption that sexual desire is too primitive to be an emotion, or that emotions cannot involve principles that prescribe pursuing certain ends, then some Kantian desires are not emotions; for Kant speaks of sexual inclinations (*Anth.*7:268), and, as mentioned in Chapter 3, he says that passion "always presupposes a maxim on the part of the subject, to act according to an end prescribed to him by his inclination." (*Anth.* 7:266). By adding such extra-textual premises about particular emotional terms, or about the nature of emotions, we could narrow down the realm of "Kantian" emotions; and we could then look for what is common to all the remaining mental states in order to provide the necessary *and* sufficient conditions on being a "Kantian" emotion. "Kantian" here is in scare quotes, however, because the resultant account would be

partially based not on anything Kant says, but rather on a certain conception of emotions that is not necessarily Kant's.

We should be wary, however, of the temptation to recruit Kant to one of the camps participating in the contemporary debate on the nature of emotions on the basis of merely-superficial similarities. It might be tempting, for example, to think that if, as I have argued, Kantian emotions, of both the feeling and the desire kind, involve *feelings* (of pleasure or displeasure), then Kant has a “feeling theory” of emotion. But feeling theories of emotion differ in their conceptions of the feelings which emotions consist in, and some of these conceptions are different from Kant's. For example, on James' understanding, the feelings that emotions consist in are perceptions of bodily changes (1884). In contrast, none of Kant's characterizations of feelings - as subjective representations (*MM* 6:211; *CPJ* 5:189, 206), as evaluations (*CPrR* 5:9 fn., *LM* 28:894), and as grounds of certain behaviors (*CPJ* 5:220, *FI* 20:230-2) - mentions bodily changes as essential to them; and if what I have argued is correct, feelings represent agreement or disagreement with one's ability or intention to act (and not bodily changes). So, Kant does not have a feeling theory of emotion in James' sense.¹³⁹ In order to avoid equivocating on the meaning of “feeling”, we would have to first spell out just what is meant by a “feeling theory” of emotion, and then see whether Kant has a set of mental states conceived in a similar manner. But even if Kant did have such a set - for instance, a set of mental states conceived as perceptions of

¹³⁹ Cohen argues that Kant's position, at least with regards to mental feelings, is the reverse of James' (1884) in that “many bodily sensations usually associated with emotions turn out to be physiological reactions of the body to mental feelings – they are neither reducible to them, nor their cause, but rather their effect on the body.” (Cohen 2020, 454, fn. 76). This claim seems to me correct insofar as in the passages Cohen mentions as supporting her claim, Kant describes certain bodily changes as caused by feelings. In one of these passages, Kant says that quickly rising hope and fear, later labeled as affects (*Anth.* 7:254), can arouse sensations of warm and cold (*Anth.* 7:154, here “sensations” probably means “objective sensations”, i.e., ones that can contribute to cognition); in another, he says that “the affect of sadness produces the physical agitation of sighing, and the affect of fright produces a scream.” (*LA* 25:600).

bodily changes - this would allow us to conclude that Kant has a “feeling theory” of emotion only by adding the further assumption that such a theory is *correct*. It is this assumption that would allow us to conclude that the set of mental states in Kant, construed in a manner similar to our assumed theory, is a set of emotions, and that Kant's theory of this set is a theory of emotion. The strategy I've proposed, in contrast, can be endorsed independently of our preferred theory of emotion.

Even without adding such extra-textual assumptions about emotions, however, I think the work done here is useful in that we can accept the accounts it provides of desires and feelings in Kant even without accepting them as accounts of the necessary conditions on being a Kantian emotion. And if we do accept them as such, I think the resultant picture is illuminating insofar as it tells us that all Kantian emotions involve pleasures or displeasures, and as such they involve a phenomenological, evaluative, and motivational component. In 2.4, I pointed out that these are the three functions traditionally associated with pleasure, and that Kant's account is attractive in that it can explain how they all hang together in one mental state. But the same three functions have also traditionally been assigned to emotion, and theories of emotion may be classified in accordance with the aspect which they take to be most essential. If my account is correct, Kant's theory of emotions of the feeling kind takes all three aspects to be essential to these mental states, without prioritizing one component over the others; in this respect, Kant's theory anticipates recent hybrid models of emotions (e.g., Lazarus 1991, Helm 2009).¹⁴⁰ His theory of

¹⁴⁰ Kant's theory of emotions of the feeling kind differs from Lazarus', however, in that Lazarus takes the appraisal component to be the cause of emotion's other components (i.e., action tendency, physiological change, and subjective affect) (Lazarus 1991, 819). In contrast, on Kant's view as I presented it in 2.4, the three components of feeling are different aspects of the same mental state.

emotions of the desire kind, in contrast, prioritizes the motivational component as the most essential, and so it resembles motivational theories of emotion.¹⁴¹

If this is correct, then the “pain model” on Cohen's list in the Introduction cannot be right. Emotions of the feeling kind are not brute sensations, if by this is meant that they have no representational content; and neither they nor emotions of the desire kind are independent of reason, if by this is meant that reason can have no impact on them. For although emotions of the feeling kind are sometimes referred to as “sensations”, they (and all emotions, by virtue of involving them) involve evaluations of objects' agreement or disagreement with the subject's ability or intention to act. These evaluations are not cognitions, as they cannot inform us about objective properties, but they are also not independent of reason. They rather provide us with important information pertinent to practical deliberation, and as such they function as defeasible reasons for action. Moreover, although some emotions involve or consist in feelings that arise from sense perceptions and as such may be said to be disconnected from our higher mental faculties, others involve feelings caused by representations of the higher faculties (such as respect for the moral law, caused by reason, or pure aesthetic pleasure, caused by the free play of the imagination and understanding). And our examination of the kind of treatment Kant prescribes for feelings and inclinations has shown that although we cannot be constrained to have or experience feelings and inclinations, we are nevertheless responsible for cultivating feelings that are not affects and for disciplining inclinations.

¹⁴¹ There are, however, also important differences between Kant's account of emotions of the desire kind and contemporary motivational theories of emotion. In contrast to Deonna and Teroni (2012), who take emotions to consist in felt bodily stances of action readiness, Kant does not take the physiological component to be essential to emotions (and although feelings are *involved* in all emotions, on Kant's view, they are not constitutive of desires). In contrast with Scarantino, who takes emotions to be causes of states of action readiness (2014), Kant identifies emotions of the desire kind with dispositions to actions, not with the causes of these dispositions.

Like Deimling's model (2014), on my account some feelings and some desires qualify as emotions. But in contrast with Deimling, my account leaves it open just which feelings and desires are emotions. Like Cohen's (2020), my account takes emotions to involve both a phenomenological (“affective”, in Cohen's terms) and an evaluative (“appraisal”) component; but while Cohen thinks motivation is not essential to feelings, I have argued that although feelings do not motivate immediately, both desires and feelings dispose us to certain behaviors (i.e., to preserve or change our state, and to produce or not to produce objects). And my account agrees with Frierson's (2014b), insofar as Kant does not have a single theory of emotion. But while Frierson thinks the domain of emotions includes products of different mental faculties, and that some emotions are products of more than one faculty, on my account each emotion - although it may be causally and intentionally connected with mental states of the other faculties - is the product of a single faculty (either the faculty of desire or feeling).¹⁴²

One worry one might have regarding the resultant account is that it splits the emotional domain into two essentially different sets of mental states, i.e., feelings, which originate in the feeling of pleasure and displeasure, and desires, which originate in the faculty of desire. But emotions as we understand them today are usually regarded as a set of mental states that have a single common nature (a natural kind), or at least share some common properties that can be captured by scientific or philosophical generalizations (a theoretical kind). And this might lead us to prefer an account which identifies Kantian emotions with a set of mental states all of which are associated with the same mental faculty.

¹⁴² My account also agrees with DeWitt's (2018) insofar as we both take emotions to involve an evaluative and a motivational component; but DeWitt takes phenomenology to be inessential to emotions, while I think it is essential to emotions of the feeling kind. Our accounts also differ in our understanding of the evaluative component, and in that DeWitt seems to want to exclude desires from the domain of emotions (although she says that inclinations and predispositions involve judgments analogous to those which are constitutive of feelings, DeWitt 2018, 81).

In response to this worry it is important to first note that even on Cohen's account of emotions as feelings, inclinations, desires, instincts and passions count as emotions; and so, her view appears to be that Kantian emotions include mental states associated with both the faculty of feeling and the faculty of desire. An alternative account on which all Kantian emotions are of the same natural or theoretical kind would have to identify them with a set of mental states of a single mental faculty (either feelings or desires),¹⁴³ or deny that the fact that different mental states are associated with different faculties entails that they are not of the same natural or theoretical kind. But as we have seen, there is evidence that some of the mental states we call emotions are associated by Kant with the faculty of feeling, while others are associated with the faculty of desire. And since Kant's view appears to be that the mental states of each faculty have a designated unique function, it seems that if some feelings and some desires are emotions, emotions come in two natural or theoretical kinds. This, however, need not count against our reconstruction, for two reasons. First, because as noted in Chapter 1, some theoreticians deny that folk emotion categories designate natural kinds (Rorty 2004; Griffiths 1997; Kagan 2007). If they are right, it is a virtue rather than a vice of Kant's account that it regards both feelings and desires as emotions. Second, and also as noted in Chapter 1, a division within the domain we nowadays call "emotional" into two theoretical categories was customary among philosophers prior to the 19th century (Dixon 2012, 339). This does not speak in favor of the philosophical tenability of the account I attribute to Kant; but it does lend some credence to its historical accuracy.

¹⁴³ This is how Williamson seems to proceed, as she identifies Kantian emotions with "intellectually caused feelings" (Williamson 2014, 2), and criticize commentators who count Kant's remarks on passion as part of his theory of emotion. But it is not clear whether she wants to deny that mental states of the faculty of desire can be emotions, or rather to propose an account of emotions in which feelings rather than desires take center stage.

Finally, if we insist that emotions must be of one natural or theoretical kind (and agree that some feelings and some desires in Kant are emotions), we have seen in Chapter 3 that there are some passages where Kant speaks of desire as a kind of satisfaction, feeling or pleasure. These passages may be used to reconstruct the account of emotions Kant should have had, had he shared the belief that emotions are of a single theoretical or natural kind. But such a reconstructed account would have to explain why, despite being assigned to two different faculties, and despite their apparent differences (in temporal directness and their relation to action), feelings and desires are nonetheless two species of one natural or theoretical kind. Rather than engaging in this exercise of forging a Kantian reconstruction that would conform with contemporary one-kind theories of emotion, I think it is best to take Kant's two-kinds account for what it is: a viable and rich alternative.

Note on Translations and Abbreviations

Kant's works are cited according to the *Akademie* edition's volume and page number, separated by a colon, and most quotations are given from the translations in the *Cambridge Edition of the Works of Immanuel Kant*. Three exceptions are the following: (i) references to the *Critique of Pure Reason* are according to the standard "A/B" format; (ii) I have provided my own translations of passages that have not yet appeared in the *Cambridge* edition; (iii) the translation of German phrases within square brackets has been modified. In addition, German words in normal brackets that appear in citations are taken for the *Akademie* edition, but do not appear in the *Cambridge* edition. Kant's *Reflexionen* are cited using Adickes' numberings in addition to the *Akademie* edition's page numbers.

I use the following abbreviations for citing Kant's works:

Anth. *Anthropologie in pragmatischer Hinsicht = Anthropology from a Pragmatic Point of View*. 2007. In *Anthropology, History, and Education*, translated by Robert B. Loudon, edited by Günter Zöllner and Robert B. Loudon, 231-429. Cambridge: Cambridge University Press.

CPrR *Kritik der praktischen Vernunft = Critique of Practical Reason*. 1996. In *Practical Philosophy*, translated and edited by Mary J. Gregor, general introduction by Allen W. Wood, 139-271. Cambridge: Cambridge University Press.

XXXA /XXXB *Kritik der reinen Vernunft = Critique of Pure Reason*. 1998. Translated and edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

- CPJ* *Kritik der Urteilkraft = Critique of the Power of Judgment*. 2000. Translated by Paul Guyer and Eric Matthews, edited by Paul Guyer. Cambridge: Cambridge University Press.
- FI* *Erste Einleitung in die Kritik der Urteilkraft = First Introduction to the Critique of the Power of Judgment*. 2000. In *Critique of the Power of Judgment*, translated by Paul Guyer and Eric Matthews, edited by Paul Guyer, 4-51. Cambridge: Cambridge University Press.
- GMM* *Grundlegung zur Metaphysik der Sitten = Groundwork of the Metaphysics of Morals*. 1996. In *Practical Philosophy*, translated and edited by Mary J. Gregor, general introduction by Allen W. Wood, 37-108. Cambridge: Cambridge University Press.
- Inquiry* *Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral = Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality*. 1992. In *Theoretical Philosophy, 1755-1770*, translated and edited by David Walford and Ralf Meerbote, 243-286. Cambridge: Cambridge University Press.
- LA* *Lectures on Anthropology*. 2012. Translated by Robert R. Clewis, Robert B. Loudon, G. Felicitas Munzel, and Allen W. Wood. Edited by Allen W. Wood and Robert B. Loudon. Cambridge: Cambridge University Press (1996).
- LE* *Lectures on Ethics*. 1997. Translated by Peter Heath, edited by Peter Heath and J.B. Schneewind. Cambridge: Cambridge University Press.
- LM* *Lectures on Metaphysics*. 1997. Translated and edited by Karl Ameriks and Steve Naragon. Cambridge: Cambridge University Press.

- LP* *Pädagogik = Lectures on Pedagogy*. 2007. In *Anthropology, History, and Education*, translated by Robert B. Loudon, edited by Günter Zöllner and Robert B. Loudon, 437-485. Cambridge: Cambridge University Press.
- LR* *Lectures on the Philosophical Doctrine of Religion*. 1996. In *Religion and Rational Theology*, translated by Allen W. Wood, edited by Allen W. Wood and George Di Giovanni, 341–451, Cambridge: Cambridge University Press.
- Mensch.* *Immanuel Kant's Menschenkunde oder philosophische Anthropologie. Nach handschriftlichen Vorlesungen herausgegeben von Fr. Ch. Starke*. 1976. Leipzig: Olms (1831). Reprint: Hidelrshim (1976).
- MM* *Die Metaphysik der Sitten = The Metaphysics of Morals*. 1996. In *Practical Philosophy*, translated and edited by Mary J. Gregor, general introduction by Allen W. Wood, 365–603. Cambridge: Cambridge University Press.
- R.* *Reflexionen aus dem handschriftlichen Nachlaß = Notes and Fragments*. 2005. Translated by Curtis Bowman, Paul Guyer, and Frederick Rauscher. Edited by Paul Guyer. Cambridge: Cambridge University Press.
- Rel.* *Die Religion innerhalb der Grenzen der bloßen Vernunft = Religion within the boundaries of mere reason*. 1996. In *Religion and Rational Theology*, translated by George Di Giovanni, edited by Allen W. Wood and George Di Giovanni, 57–215. Cambridge: Cambridge University Press.
- SF* *Der Streit der Fakultäten = The Conflict of the Faculties*. 1996. In *Religion and Rational Theology*, translated by Mary J. Gregor and Robert Anchor, edited by Allen W. Wood and George Di Giovanni, 233–327. Cambridge: Cambridge University Press.

Bibliography

- Allison, Henry E. 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- . 1998. "Pleasure and Harmony in Kant's Theory of Taste: A Critique of the Causal Reading." In *Kants Ästhetik / Kant's Aesthetics / L'esthétique de Kant*, edited by Herman Parret, 466–83. Berlin: de Gruyter.
- . 2020. *Kant's Conception of Freedom: A Developmental and Critical Analysis*. Cambridge: Cambridge University Press.
- Baron, Marcia W. 1995. *Kantian Ethics Almost without Apology*. Ithaca: Cornell University Press.
- Baum, Manfred. 2006. "Gefühl, Begehren und Wollen in Kants praktischer Philosophie." *Jahrbuch für Recht und Ethik / Annual Review of Law and Ethics*, 14:125-39.
- Baxley, Anne Margaret. 2010. *Kant's Theory of Virtue: The Value of Autocracy*. New York: Cambridge University Press.
- Borges, Maria. 2004. "What Can Kant Teach Us about Emotions?" *The Journal of Philosophy* 101 (3): 140-58.
- . 2008. "Physiology and the Controlling of Affects in Kant's Philosophy." *Kantian Review* 13 (2): 46–66.
- . 2019. *Emotion, Reason and Action in Kant*. London: Bloomsbury Academic.
- Byrne, Alex. 2006. "Intentionality." In *The Philosophy of Science: An Encyclopedia*, edited by Sahotra Sarkar and Jessica Pfeifer, 406-10. London: Routledge.
- Clewis, Robert R. 2009. *The Kantian Sublime and the Revelation of Freedom*. Cambridge: Cambridge University Press.

- Cohen, Alix. 2017. "Kant on Emotions, Feelings, and Affectivity." In *The Palgrave Kant Handbook*, edited by Matthew C. Altman, 665-681. London: Palgrave Macmillan.
- . 2020. "A Kantian Account of Emotions as Feelings." *Mind* 129 (514): 429-60.
- Craig, A.D. 2002. "How do you feel? Interoception: the sense of the physiological condition of the body." *Nature Reviews Neuroscience*, 3: 655–66.
- . 2009. "How do you feel – now? The anterior insula and human awareness." *Nature Reviews Neuroscience*, 10 (1): 1059–70.
- . 2014. *How Do You Feel? An Interoceptive Moment with Your Neurobiological Self*. Princeton: Princeton University Press.
- Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Deimling, Wiebke. 2014. "Kant's Pragmatic Concept of Emotions." In *Kant on Emotion and Value*, edited by Alix Cohen, 108-25. London: Palgrave Macmillan.
- . 2018. "Two Different Kinds of Value? Kant on Feeling and Moral Cognition." In *Kant and the Faculty of Feeling*, edited by Kelly Sorensen and Diane Williamson, 25-40. Cambridge: Cambridge University Press.
- Deonna, Julien A. and Fabrice Teroni. 2012. *The Emotions: A Philosophical Introduction*. London: Routledge.
- DeWitt, Janelle. 2014. "Respect for the Moral Law: The Emotional Side of Reason." *Philosophy* 89 (1): 31-62.
- . 2018. "Feeling and Inclination: Rationalizing the Animal Within." In *Kant and the Faculty of Feeling*, edited by Kelly Sorensen and Diane Williamson, 67-87. Cambridge: Cambridge University Press.

- Dixon, Thomas. 2012. "'Emotion': the History of a Keyword in Crisis." *Emotion Review*, 4 (4): 338–44.
- Duncker, Karl. 1941. "On Pleasure, Emotion, and Striving." *Philosophy and Phenomenological Research* 1 (4), 391-430.
- Engstrom, Stephen. 2010. "Reason, desire, and the will." In *Kant's Metaphysics of Morals: A Critical Guide*, edited by Lara Denis, 28-50. Cambridge: Cambridge University Press.
- Fahmy, Melissa Seymour. 2009. "Active Sympathetic Participation: Reconsidering Kant's Duty of Sympathy." *Kantian Review*, 14 (1): 31–52.
- Frierson, Patrick, R. 2014a. "Affective Normativity." In *Kant on Emotion and Value*, edited by Alix Cohen, 166-190. London: Palgrave Macmillan.
- . 2014b. *Kant's Empirical Psychology*. Cambridge: Cambridge University Press.
- Geiger, Ido. 2011. "Rational Feelings and Moral Agency." *Kantian Review* 16 (2): 283–308.
- Ginsborg, Hannah. 2003. "Aesthetic Judging and the Intentionality of Pleasure." *Inquiry* 46 (2): 164–81.
- Giordanetti, Piero. 2015. "Begierde." In *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr, and Stefano Bacin, 233. Berlin: De Gruyter.
- Goldie, Peter. 2000. *The Emotions: A Philosophical Explanation*. Oxford: Oxford University Press.
- Grenberg, Jeanine M. 2001. "Feeling, Desire and Interest in Kant's Theory of Action." *Kant-Studien* 92 (2): 153-79.
- Griffiths, Paul E. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.

- Guyer, Paul. 1997. *Kant and the Claims of Taste*, second edition. Cambridge: Cambridge University Press.
- . 2005. *Kant's System of Nature and Freedom: Selected Essays*. Oxford: Clarendon Press.
- . 2010. "Moral Feelings in the Metaphysics of Morals." In *Kant's Metaphysics of Morals: A Critical Guide*, edited by Lara Denis, 130-51. Cambridge: Cambridge University Press.
- . 2018. "What is it Like to Experience the Beautiful and Sublime?" In *Kant and the Faculty of Feeling*, edited by Kelly Sorensen and Diane Williamson, 147-65. Cambridge University Press.
- Helm, Bennet W. 2002. "Felt Evaluations: A Theory of Pleasure and Pain." *American Philosophical Quarterly* 39 (1): 13-30.
- . 2009. "Emotions as Evaluative Feelings." *Emotion Review* 1 (3): 248-55.
- Herman, Barbara. 1993. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- Höwing, Thomas. 2013. *Praktische Lust: Kant über das Verhältnis von Fühlen, Begehren, und Praktische Vernunft*. Berlin: De Gruyter.
- James, William. 1884. "What is an Emotion?" *Mind* 9 (34): 188–205.
- Kagan, Jerome. 2007. *What is Emotion? History, Measures, and Meanings*. New Haven, CT: Yale University Press.
- Kain, Patrick. 2003. "Prudential Reason In Kant's Anthropology." In *Essays on Kant's Anthropology*, edited by Brain Jacobs and Patrick Kain, 230-65, Cambridge: Cambridge University Press.
- Katz, Leonard D. 2016. "Pleasure." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2016/entries/pleasure/>](https://plato.stanford.edu/archives/win2016/entries/pleasure/).

- Kulstad, Mark, and Laurence Carlin. 2020. "Leibniz's Philosophy of Mind." *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/win2020/entries/leibniz-mind/>.
- Lazarus, Richard S. 1991. "Progress on a Cognitive-Motivational-Relational Theory of Emotion." *American Psychologist* 46 (8): 819–34.
- Lewis, David. 1972. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (3): 249–58.
- Locke, John. 1975. *An Essay Concerning Human Understanding*. Edited by Peter H. Nidditch. Oxford: Clarendon Press.
- Louden, Robert B. 2002. *Kant's Impure Ethics: From Rational Beings to Human Beings*. Oxford: Oxford University Press.
- Messina, James, and Sturm, Thomas. 2015. "Erklärung." In *Kant-Lexikon*, edited by Marcus Willaschek, Jürgen Stolzenberg, Georg Mohr, and Stefano Bacin, 559-60. Berlin: De Gruyter.
- Morrisson, Iain P.D. 2008. *Kant and the Role of Pleasure in Moral Action*. Ohio: Ohio University Press.
- Munzel, G. Felicitas. 1999. *Kant's Conception of Moral Character: The 'Critical' Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press.
- Newton, Alexandra. 2017. "Kant On Animals And Human Pleasure." *Canadian Journal of Philosophy* 47 (4): 518-40.
- Nussbaum, Martha C. 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Papish, Laura. 2007. "The Cultivation of Sensibility in Kant's Moral Philosophy." *Kantian Review* 12 (2): 128–46.

Rorty, Amélie Oksenberg. 2004. "Enough Already with 'Theories of the Emotions'." In *Thinking About Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. Solomon, 269-78. New York: Oxford University Press.

Rueger, Alexander. 2020. "Kant on Feelings, Sensations, and the Gap Between Rationality and Morality." *Kantian Review* 25 (1): 125-48.

Rosenkranz, Melissa A., Daren C. Jackson, Kim M. Dalton, Isa Dolski, Carol D. Ryff, Burt H. Singer, Daniel Muller, Ned H. Kalin, and Richard J. Davidson. 2003. "Affective style and *in vivo* immune response: Neurobehavioral mechanisms." *Proceedings of the National Academy of Sciences of the U.S.A.* 100 (19): 11148–52.

Ryle, Gilbert. 1949. *The Concept of Mind*. London: Hutchinson.

Sabini, John, and Maury Silver. 1987. "Emotions, Responsibility, and Character." In *Responsibility, Character and the Emotions*, edited by Ferdinand Schoeman, 165-75. Cambridge: Cambridge University Press.

Scarantino, Andrea. 2014. "The Motivational Theory of Emotions." In *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*, edited by Justin D'Arms and Daniel Jacobson, 156-85. Oxford: Oxford University Press.

---. 2016. "The Philosophy of Emotions and Its Impact on Affective Science." In *Handbook of Emotions*, fourth edition, edited by Lisa Feldman Barrett, Michael Lewis, and Jeannette M. Haviland-Jones, 3-48. New York, NY: Guilford Press.

Schönecker, Dieter. 2010. "Kant über Menschenliebe als moralische Gemütsanlage." *Archiv für Geschichte der Philosophie* 92 (2): 133-75.

- Schroeder, Tim. 2020. "Desire." *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/sum2020/entries/desire/>](https://plato.stanford.edu/archives/sum2020/entries/desire/).
- Schmitter, Amy M. 2010. "17th and 18th Century Theories of Emotion." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2016/entries/emotions-17th18th/>](https://plato.stanford.edu/archives/win2016/entries/emotions-17th18th/).
- Sherman, Nancy. 1997. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press.
- Solomon, Robert C. 1980. "Emotions and Choice." In *Explaining Emotions*, edited by Rorty, Amélie Oksenberg, 251–81. Los Angeles: University of California Press.
- Sorensen, Kelly D. 2002. "Kant's Taxonomy of the Emotions." *Kantian Review* 6 (1): 109-28.
- Thayer, Robert E. 1989. *The Biopsychology of Mood and Arousal*. New York: Oxford University Press.
- Thomason, Krista K. 2017. "A Good Enough Heart: Kant and the Cultivation of Emotions." *Kantian Review* 22 (3):441-62.
- Williamson, Diane. 2015. *Kant's Theory of Emotion: Emotional Universalism*. London: Palgrave Macmillan.
- Wilson, Eric E. 2016. "Habitual Desire: On Kant's Concept of Inclination." *Kantian Review* 21 (2): 211-235.
- Wood, Allen, W. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.
- . 2003. "Kant and the Problem of Human Nature." In *Essays on Kant's Anthropology*, edited by Brian Jacobs and Patrick Kain, 38-59. Cambridge: Cambridge University Press.

---. 2014. "Empirical Desire." In *Kant's Lectures on Anthropology: A Critical Guide*, edited by Alix Cohen, 133-50. Cambridge: Cambridge University Press.

Wuerth, Julian. 2014. *Kant on Mind, Action and Ethics*. Oxford: Oxford University Press.

Zuckert, Rachel. 2002. "A New Look at Kant's Theory of Pleasure." *The Journal of Aesthetics and Art Criticism* 60 (3): 239-252.

Curriculum Vitae

Education

B.A. in Behavioral Sciences, 2008, Ben Gurion University, Israel, *cum laude*.

M.A. in Philosophy, 2012, Tel Aviv University, Israel, *magna cum laude*. Dissertation: *The Non-Moral Conscience in Kant and Heidegger*. Supervised by Ruth Ronen.

Ph.D. in Philosophy, minor in History of German Thought, 2021, Indiana University, Bloomington. Dissertation: *Kant's Theory of Emotion: Toward a Systematic Reconstruction*.

Chair: Allen Wood. Committee: Marcia Baron, Sandra Shapshay, Johannes Türk.

Research

Areas of Specialization: Kant, Kant's ethics, Kant's moral psychology.

Areas of Competence: Ethics, Moral Psychology, Philosophy of Emotions, Early Modern Philosophy.

Papers

2000. "Which Emotions Should Kantians Cultivate (and Which Ones Should They Discipline)?" *Kantian Review* 25 (I): 53-76.

Work in Progress

"Kantian Desires: A Holistic Account." Conditionally accepted.

"Are Kantian Emotions Feelings?" Under "revise and resubmit."

Presentations

May 2014, 17th Annual Meeting of the Israeli Philosophy Association. Title: "Leaving Akrasia Behind: Kant on Weakness of Will."

April 2016, Eastern Study Group of the North American Kant Society Meeting at Yale University. Title: "The Trouble with Happiness: Kant's Purification of the Moral Incentive."

June 2016, Leuven Kant Conference. Title: "The Trouble with Happiness: Kant's Purification of the Moral Incentive."

March 2018, Universität Siegen (Schönecker Colloquium). Title: "How Cognitively Complex Are Kantian Emotions?"

May 2018, Humboldt University (Rosefeldt Colloquium). Title: "Kantian Desires: The Cultural and the Natural."

(October 2018, Multilateral Kant Colloquium at Cantina, Italy. Title: "Kant and Emotions: The Importance of the Feeling-Desire Distinction" (**could not attend**).

April 2019, Indiana University (Dissertation Group Colloquium). Title: "Why Kant May Have a Hybrid Model of Pleasure: A Response to Guyer."

September 2019, Universität Siegen (Schönecker Colloquium). Title: "Does Kant Define Emotions as Feelings? A Response to Alix Cohen."

December 2019, FU Berlin (Emundts Colloquium). Title: "Kant's Hybrid Model of Pleasure."

May 2020, HU Berlin (Rosefeldt Colloquium). Title: "Kantian Desires: A Holistic Account."

Teaching

2009	Guided Reading in Descartes' <i>Meditations</i> (Tel Aviv University)
2010	Guided Reading in Hume's <i>Treatise</i> , Book I (Tel Aviv University)
2011	Guided Reading in Plato's <i>Gorgias</i> and Kant's <i>Groundwork</i> (Tel Aviv University)
2014	Kant's Categorical Imperative (Assistant Instructor to Prof. Ido Geiger, Ben Gurion University)
2015-16	Introduction to Philosophy (Assistant Instructor to Prof. Gary Ebbs, Indiana University)
	Introduction to Ethics (Assistant Instructor to Prof. Timothy O'Connor, Indiana University)
2016-17	Introduction to Ancient Philosophy (grader and substitute lecturer for Prof. Pieter Sjoerd Hasper, Indiana University)
	Introduction to Philosophy (Assistant Instructor to Prof. Frederick F. Schmitt, Indiana University)
2019	Introduction to Ethics (Instructor, Indiana University)
	All Assistant Instructor positions at Indiana University involved leading two stand-alone sections per week and grading.

Awards and Honors

- 2009, 2011 Tel Aviv University's "Poesis" Prize for Academic Excellence
- 2011-2013 Fellow at the Program for the Study of the Philosophy of Immanuel Kant and its Legacy (held jointly by Tel Aviv and Ben Gurion University)
- 2012 - Present Co-Editor of *Mafte'akh: Lexical Review of Political Thought*
- 2012-2014 Coordinator, book review series, Minerva Humanities Center, Tel Aviv University
- 2014 Indiana University's Philosophy Department's First Year Fellowship
- 2017 Indiana University's Irving and Shirley Brand Graduate Fellowship
Indiana University's Philosophy Department's Academic Excellence Award
- 2017 Indiana University - Freie Universität, Berlin, Foreign Exchange Fellowship
- 2019 Minerva Stiftung Doctoral Fellowship. Project Title: "What (Else) Can Kant Teach Us About Emotions?" Hosted by Prof. Dina Emundts at Freie Universität, Berlin

Languages

- English – fluent (reading, writing, and speaking)
- Hebrew – native (reading, writing, and speaking)
- German - fluent reading, intermediate writing and speaking
- French – elementary (reading)
- Farsi - intermediate (reading, writing)
- Arabic - elementary (reading, writing)

Dissertation Summary

Those of us trying to put together Kant's theory of emotion are facing a difficult task for two reasons. First, Kant has no term which is an obvious equivalent of "emotion" as used in English nowadays; he rather uses a variety of terms which look like good candidates to be included in a theory of emotion, but are associated by him with two different mental faculties. For example, feeling and affect are associated by Kant with the faculty of feeling, and inclination and passion are associated with the faculty of desire. Second, theorists disagree about the nature of emotions, construing them as feelings, motivations, evaluations, or some combination of these phenomena. Taken together, these facts make it difficult to determine where in Kant we should be looking to find out how he understands emotions. Should we focus on a single term in Kant, as Cohen does in her "Kantian account of emotions as feelings"? Should we rather look at various terms, which originate in different mental faculties, as Deimling does? Or is it rather the case that Kant has no unique model or general theory of emotion at all, as Borges and Frierson argue (respectively)?

In my dissertation, I propose an alternative to these approaches to Kant's theory of emotion, novel both in its methodology and upshot. I begin by tackling the problem of how to reconstruct Kant's theory of emotion in light of the aforementioned difficulties by examining recent attempts to do so. The available accounts vary to a significant extent, I propose, because they are based on different assumptions about emotions: that certain terms in Kant designate emotions (Cohen), that emotions track complex values and are under our mediate control (Deimling), and that certain terms are generally considered to designate emotions while others are not so considered (Frierson). Some of these assumptions are questionable or vague, which renders the reconstructions based upon them questionable too. We could, nevertheless, make headway toward an agreed upon reconstruction if we started from the generally accepted and reasonable premise that what we call "emotions" refers in Kant to a set of mental states that originate in the feeling of pleasure and displeasure and in the faculty of desire, and then proceeded to look at what Kant says about the nature of these two kinds of mental states.

In accordance with this plan, I start by looking at Kant's characterizations of the mental states that originate in the feeling of pleasure and displeasure, i.e., "feelings" or "pleasure" and "displeasure." Some commentators argue that they are phenomenological qualities, identical within their "positive" and "negative" tokens (Guyer); others maintain that they combine judgments and dispositions to action (DeWitt); still others stress their dispositional nature (Guyer in his more recent work, and Rueger). Rather than choosing between them, I propose a way of combining what each of these accounts gets right. On my proposal, pleasure brings together three phenomena: a felt phenomenological quality (that need not feel the same across its tokens), an evaluation of an object's agreement with one's ability or intention to act, and a disposition to preserve one's state or change it, or to act in order to produce or not to produce an object.

I proceed to examine Kant's characterizations of the mental states which originate in the faculty of desire, i.e., desires (and aversions). In contrast with Frierson, who stresses desire's causal connection with action, I argue that desires dispose to action, but need not bring it about; and in contrast with readings that stress desire's connection with feeling (Wood, Grenberg), I argue that desires (other than wishes) are directed at future objects, and so they differ from the feelings they are causally connected to, which refer to the subject's affective state. The proposed account has the merit of being compatible with Kant's talk of desires that do not lead to action, and of explaining the difference between desires and the feelings they are causally connected to.

The final chapter makes use of my account of feelings and desires to examine the kind of treatment Kant prescribes for them. Against Thomason's claim that Kant does not think we should cultivate emotions, I argue that Kant's view is that feelings (except affects) should be cultivated, while inclinations - although they can and perhaps may be cultivated - should generally be disciplined. The right treatment for an emotion depends on its ability to motivate action immediately, and on the threat it poses for rational deliberation.

These chapters provide the foundation for a rich and complex account, on which Kantian emotions come in two kinds: feelings, and desires. Emotions that are feelings are evaluations of objects' relations to the subject, that have a felt quality and dispose us to action by means of desires; emotions that are desires are dispositions to action that need not bring it about, and are usually directed at future objects. Although it is possible that this account provides the necessary but insufficient conditions on a mental state's being a Kantian emotion, it nevertheless has the merit that it rests on explicit, reasonable and generally accepted assumptions. Moreover, it provides a compelling alternative to contemporary theories of emotion, that combines the three phenomena traditionally associated with them.