Introduction

This work presents the first systematic, exegetical, and comprehensive study of the concept of love in the philosophy of Immanuel Kant. Love is often considered to be among the most important yet perplexing of all human phenomena, and Kant is generally thought to be among the greatest philosophers in the Western tradition. I thus find it remarkable that Kant’s views on love have not previously been investigated in much depth, or with an outlook on the way the concept of love figures and operates within his philosophy as a whole.

My research will show that love is actually important to Kant’s philosophy, at any rate a lot more important than commonly assumed. It may come as a surprise to see how often Kant thinks about love, how much he writes about it, and that he holds various philosophical views about it. In particular, an understanding of how the concept of love functions in Kant’s practical philosophy is necessary for an overall understanding of his ethical project. Even though we might think otherwise at first glance, love plays an integral role in Kant’s conception of human life.

Love is a complex and multifaceted concept, and Kant’s philosophical ideas on love yield no exception to this fact. Not only are these ideas important for understanding Kant, I also think that his views on love are interesting as such. What the views are and how they fit together is the topic of the present volume. If one is generally open to learning from Kant, this book will hopefully show that it is also possible to learn from what he thought about love. The nature of my work is mostly interpretative, and in particular I will formulate two exegetical claims, which mark the core of the study. First, I hold that in Kant we can detect a general division of love, according to which love in general divides into love of benevolence [Liebe des Wohlwollens] and love of delight [Liebe des Wohlgefallens]. The general division of love in Kant is a key for understanding love in Kant. Second, I hold that by identifying various aspects of love in Kant, such as self-love, sexual love, love of God, love of neighbour, and love in friendship, and by studying the various things he says about the different aspects of love, we can detect an ascent of love from the strongest impulses of human nature to the highest ideals of morally deserved happiness. It is these two claims that will be clarified and defended during the course of the work.
II

There are of course reasons for the lacuna in previous research. In general, the study of Kant’s moral philosophy or ethics, broadly construed, has tended to emphasise the foundations of moral philosophy as articulated in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*. Traditionally, fewer resources have been directed to research on Kant’s last major contribution to moral philosophy, *The Metaphysics of Morals*, where love figures much more prominently than in the *Groundwork* or the *Critique of Practical Reason*. In the anglophone research community, the tendency to give less weight to *The Metaphysics of Morals* can also be at least partially explained by the fact that a reliable, complete translation of the entire work has only been made available in recent decades. The emphasis on the *Groundwork* in relation to *The Metaphysics of Morals* has tended to yield a picture of a ‘cold’ Kant, a picture of a philosopher who emphasises duty over everything else and is wary of, if not outright hostile towards, emotions or affective dispositions as part of the moral life.

When I have discussed my project with non-Kantian academic philosophers, I have often been met with surprise: ‘*Love* in Kant? Oh, I didn’t know that Kant said anything at all about love! If he did, surely it must have been some kind of antithesis of love that he was really after...’ And so on. Within the community of academic philosophers in general, Kant’s views on love are clearly not well known. Within the circle of scholars working on Kant’s ethics, however, this kind of picture no longer obtains, and as more attention has been given to the ‘Doctrine of Virtue’ in *The Metaphysics of Morals* and to Kant’s anthropological works, we now have a much fuller, more balanced, and more comprehensive picture of the emotional life of the Kantian moral agent. We now know that the role of the ‘natural’, sensory-aesthetic part of the human cognitive apparatus as related to ‘pure practical reason’ or ‘freedom’ (to use Kant’s own dualism) is much more complex and nuanced than commonly assumed by those not all that familiar with Kant’s ethics. Thanks to the work of Kantian ethicists emphasising or defending the importance of emotive elements in Kant’s moral philosophy, what we have been witnessing in the last thirty years or so is the emergence of a ‘nicer’ Kant, whose overall take on moral philosophy is ‘kinder’, ‘warmer’, or more humane than what reading merely the *Groundwork* might imply.

My work represents this ‘warmer’ school of Kantian ethics. However, even my kind of approach might appear too ‘cold’ for those who wish to ground ethics or moral norms in emotive dispositions, pain and pleasure, empathy or ‘warm-heartedness’, instinctive benevolence, or the feeling of love. The account of love in Kant that I seek is meant to be true to the foundations of Kant’s moral
thought and to the letter of his text, so that the picture I provide will be not only defensible but also exegetically balanced and accurate. For instance, even in the context of the ‘nice’ Kant presented here, the feeling of love can never be the objective foundation of morality. The foundation is pure practical reason and respect for the moral law.

I won’t be talking much about what Kant could or should have thought about the problem of love; rather, I will systematically reconstruct the positions he did hold. My aim is to arrive at a general outline of the concept of love as it figures in my chosen target system, or in other words, in the propositional natural language ‘data set’ I wish to analyse. That is, I will be investigating how the word ‘love’ [Liebe] takes on different functions and meanings when it appears in various contexts within the philosophical writings of Kant.

This kind of approach has its limitations, to be sure, and I am not out to claim that Kant’s conceptions of love simply sprang from pure reason in a historical vacuum. In the tradition of Western philosophy, investigations into love date back to the Presocratic Empedocles. In particular, the notion of Eros in Plato’s Symposium has had an immense effect on subsequent European philosophy and culture. The connection between the concept of love and the notion of the highest good, which we will continue to see in Kant, originates from Plato. Besides Plato, another decisive factor contributing to how it was possible for someone to conceptualise love in 18th-century Prussia is the Christian religion, and especially the teachings of Jesus as they were preserved in the New Testament. It is not much of an exaggeration to assert that from a decidedly historical perspective, one cannot understand the context in which Kant writes about love without acknowledging the existence of at least two historical documents: the Symposium and the Sermon on the Mount.

It would be possible, and highly interesting, I think, to trace the historical genealogy of the concept of love from Plato to Kant, to consider the parallels and continuities between, for instance, the way love figures in Aristotle’s and Kant’s conceptions of friendship, to analyse the extent to which Kant’s conceptualisations of love are indebted to, say, Augustine or Aquinas, or to his more immediate predecessors like the British moralists, or Leibniz, Wolff, or Baumgarten, or indeed to his Pietist upbringing. I would be especially inclined to point out (and this may be obvious to some) that the link between love and the highest good is common to both Platonism and Christianity, that this link is sustained through the Middle Ages in the Scholastic fusion of Plato, Aristotle, and Jesus, that it remains clearly visible in the British sentimentalists like Hutcheson, and that it influences Kant’s construction of the regulative ideality of his ethical system as a whole. But this kind of historical, cultural and comparative approach is beyond the scope of my study.
The concept of love is scattered throughout Kant’s massive corpus, and I will have enough work on my hands in just getting the exegesis right. Likewise, I will not enter the ‘post-Kantian’ domain and interpret my results in the light of later strands of German idealism; nor will I use my findings to engage with contemporary discussions in the philosophy of love. Although I believe that the approach and the general framework developed in this book can later be reworked and used to formulate a contemporary, post-Kantian philosophy of love, this is not my aim here. This study is very much Kant immanent.

As I mentioned above, Kant’s conception of love has not been investigated comprehensively or in detail in the literature to date. This is not to say that there aren’t any discussions of this topic, however. Since a key component of my argument concerns the novelty of my claims, I feel obliged to say at least something about how love has previously been analysed and discussed in the study of Kant. Therefore, I will now offer some general reflections on the state of previous research, and in doing this, I will refer to individual accounts only insofar as they are particularly representative of the points I wish to make.

The Kant-Bibliographie 1945–1990 (Malter/Ruffing 1999), which aims to present a comprehensive bibliography of Kant-related academic writings from the period between the aforementioned years, lists some 9000 titles, 5 titles of which contain the word ‘love’ or ‘Liebe’. From this one can plausibly generalise that from 1945 to 1990 the ‘love ratio’ in Kant scholarship was approximately 1/1800. Things have changed in this regard, and in the last couple of decades in particular, research on love in Kant has grown to the point where at least one or two new papers on the topic are published each year. However, the total number of research articles in the field still only amounts to a good handful. There is no danger of drowning in the secondary literature.

I think it is possible to divide the existing research roughly into three categories. First, there are those accounts that engage with a specific aspect of love or discuss love within a particular work by Kant. These accounts yield partial knowledge, and when the discussions are sufficiently detailed and well argued, they greatly promote our understanding of love in Kant. Most of the research on love in Kant belongs more or less to this category. Examples of the first kind of approach include works by Marcia Baron (2009; 2013), Melissa Fahmy (2009; 2010), and Dieter Schönecker (2010; 2013), where the focus is the ‘Doctrine of Virtue’ of The Metaphysics of Morals or neighbourly love more generally. Second, there are accounts that mention ‘love’ in the title but do not actually provide an interpretation (in any significant detail) of what Kant had to say about the issues the title refers to (e.g. Miller 1985; Vanden Auweele 2014). Third, there are accounts that at least ostensibly aim to provide a more general outlook on the concept of love in Kant, or somehow claim or attempt to articulate general prop-
positions about love in Kant (e.g. Streich 1924; Moors 2007; Schönecker 2010; Grenberg 2014). It is this third category that is actually the most interesting for the purposes of my work, since what I am after here is precisely a more general account of this kind. However, none of these previous discussions is anything close to a book-length monograph (the closest is Streich (1924)). They are not even articles dedicated solely to this issue, but rather propositions within papers, which pursue various general aims. I do not wish to come across as blaming these authors, nor am I suggesting that they should have done otherwise. What I am saying is that regarding the concept of love in Kant, the lacuna in the research is real.

Besides being generally exegetical in the sense that I interpret rather than evaluate, my approach has some other features that are worth highlighting. When I analyse the concept of love by investigating how the word ‘love’ appears in various propositional structures within a given text or set of texts, I analytically divide the concept of love into different aspects, according to how the word ‘love’ appears in different contexts. For example, a particularly high number of references to ‘love’ occur in contexts where Kant is speaking about the ‘self’ or ‘one’s own happiness’, ‘sexuality’, ‘God’, ‘neighbours’ or ‘other human beings’, and ‘friendship’. Analysed in these terms, the concept of love will consist at least of the aspects of ‘self-love’, ‘sexual love’, ‘love of God’, ‘love of neighbour’, and ‘love in friendship’. Naturally, the interrelationships between the aspects are also very important. The aspects must be organised into a whole as rationally as possible. Just how this is to be accomplished, however, is impossible to say without first becoming well acquainted with the object of study. It is by comparing the aspects with each other that a general structure of the target concept can be approached. In this way, the concept is marked out by the instances of the word in the light of the aspects that have been identified, but in contrast with the mere word, the concept comprises a more comprehensive propositional domain, which includes all the aspects (or the proposition sets the aspects consist of). The comparative arrangement of the aspects should reveal the possible regularities or invariances between aspects, so that these invariances can then be said to belong generically to the concept. As such, dividing the concept of love into different sorts of love is of course nothing new (cf. e.g. Fromm 1957; Lewis 1963), but the divisions are not self-evident, and the particular way I make these divisions is novel in the study of Kant.
To clarify my approach further, I think my investigation into the concept of love in Kant must be called *quasi-inductive*. I do not claim to be analysing every possible aspect of love, let alone every single instance of the word ‘love’ in Kant’s corpus. I am after a relatively robust yet manageable account of the concept of love, and I therefore limit the construction of the framework to those aspects that figure most prominently in Kant’s philosophy. While the method is inductive in the sense that I gather textual data and make generalisations based on sample populations of the word ‘love’, my results are neither certain nor absolutely necessary. What I will say is merely that my division of the concept of love in Kant is one possible, plausible, or viable way of setting up a framework of analysis, and that while my results remain incomplete and hopefully subject to criticism, this study is nevertheless the first comprehensive *approximation* of what the concept of love in Kant’s philosophy might look like.

To return to the basic claims made at the beginning of this introduction: what is the general division of love in Kant? As noted above, I identify five (or six) particularly important aspects of love to be discussed in the study: self-love, sexual love (and love of beauty), love of God, love of neighbour, and love in friendship. This list is not exhaustive and could be constructed otherwise, but these aspects do exist in Kant’s writings, and within this framework of aspects Kant consistently uses or implies a division between two generic kinds of love: love of benevolence [*Liebe des Wohlwollens*] and love of delight [*Liebe des Wohlgefallens*]. These two loves appear regularly. In general, love of benevolence in Kant is *goodwill that is directed to the well-being of its object*. It can be weak or active, but the wishful or actively sought end of all instances of love of benevolence is that things go well for the object, no matter how minimal one’s interest in the well-being of the object actually is. Love of delight, on the other hand, is *a pleasure taken in the physical or moral perfections, or even the sheer existence, of the object*. It does not carry an aim or an interest in the same way that love of benevolence does. Rather, it is a reaction or a response to an encounter with the object of love and its qualities. It is primarily a feeling aroused by the object in conjunction with the cognitive faculties or capacities of the agent.

In different contexts, love of benevolence and love of delight will vary according to their objects and the aspects of love to which they relate, so that they acquire somewhat different meanings and different functions depending on the aspect in question. However, I have found only one direct reference to a general division of love in Kant’s published works, and the existence of the division must be shown and systematically reconstructed with various sources for each of the individual aspects. The direct remark is contained in the first part of *Religion within the Bounds of mere Reason*, where Kant discusses the origin of evil and, more precisely, its relation to self-love. Kant refers to a general division
of love in a lengthy footnote: ‘Like love in general, self-love too can be divided into love of benevolence and love of delight (BENEVOLENTIAE ET COMPLACENTIAE), and both (as is self-evident) must be rational.’ (R, 6:45.22–25)¹ In the specific context of self-love, the general division of love basically means that we want things to go well for ourselves (love of benevolence), and that we are pleased (love of delight) when things do work out well for us (overall self-love is more complicated than this, and I discuss the complexity of the general division of self-love in ch. 1.2). The remark in the Religion also asserts that both love of benevolence and love of delight ‘must be rational’. It is not entirely clear what this means, and whether the ‘must’ [müssen] should be understood as normative or as part of the description of the loves in question. It is also not clear whether the rationality that Kant is talking about here is meant to apply to love of benevolence and love of delight generally or merely (or specifically) in the context of self-love. We know from the passage itself and from several places elsewhere that Kant allows for the existence of inclination-based love of benevolence and pathological love of delight (see e.g. ch. 4.1). It is therefore clear that the phrase ‘must be rational’ cannot be taken to mean ‘must be based on reason’. In the specific context of self-love, the rationality of love of benevolence for one-

¹ ‘Wie Liebe überhaupt, so kann auch Selbstliebe in die des Wohlwollens und des Wohlgefallens (BENEVOLENTIAE ET COMPLACENTIAE) eingetheilt werden, und beide müssen (wie sich von selbst versteht) vernünftig sein.’ It is difficult to find a satisfying translation that captures the parallelism, apparent in the German language, of the two basic forms of love. The love that is benevolence, is, at its basis, willing well. It is commonly translated as ‘good will’, ‘benevolence’, or ‘well-wishing’. Only the last, I believe, is implausible as a general strategy, since Kant often explicitly distinguishes between practical love of benevolence and mere wishing [wünschen]. The translation of Liebe des Wohlgefallens is trickier. Wohlgefallen consists of the adverbial conjunction Wohl (‘well’), the prefix ge, which implies conjoining or linking together, and the verb fallen. So literally, it signifies something like ‘fall together well’. The verb gefallen as such means to please or to delight, and Wohlgefallen is indeed sometimes translated as ‘well-pleasedness’. This is quite possible as a translation, but it suggests satisfaction in an outcome, and while Kant sometimes uses Wohlgefallen this way, there is also in Kant an underlying sense of Wohlgefallen-love as an immediate sensory impulse of positive attraction, better captured by the English word ‘delight’ than ‘well-pleased’. The English term ‘good pleasure’ has a biblical background and is also plausible, but like ‘well-pleased’ it perhaps obscures some of the immediateness and the attractive pull of Wohlgefallen. The English ‘well-liking’ (which is also used by some translators) is unwarranted, I think, for it actually means something like ‘good physical condition’. Wohlgefallen is sometimes translated as ‘satisfaction’, sometimes as ‘liking’, and other times as ‘approbation’. In the interest of remaining technically consistent, I use ‘love of benevolence’ and ‘love of delight’ in this study because I think that overall they best capture what Kant is talking about, even at the expense of the linguistic parallelism. My solution is naturally open to criticism.
self means that the inclination-based love of benevolence for oneself includes the successful long-term use of instrumental reasoning, i.e. it is rational as prudence. Rational love of delight for oneself, on the other hand, means either taking pleasure in one’s own prudence or a kind of self-contentment that is based on one’s respect for the moral law (in the last case the love of delight for oneself would be based on reason) (R, 6:45fn.; see ch. 1.2). However, if we interpret the words ‘must be rational’ as referring to love of benevolence and love of delight more generally, beyond the context of self-love, we may note that Kant never talks of mere non-rational animality in terms of the general division of love, and his usage of the terms seems to be limited to the context of rational beings. In this sense, love of benevolence and love of delight always imply reason, even when a token of benevolence or delight is pathological or based on inclination (ch. 1). This means that for Kant, love of benevolence and love of delight appear in rational creatures, entangled in their rational capacities. Some loves are derived from pure practical reason, like practical love of neighbour as the duty to be benevolent and beneficent to others (ch. 4). Similarly, love of God is an idea derived from moral reason (ch. 3). It is also the case that Kant’s ethical writings involve a demand for cultivation, which ethically means the conscious striving to make one’s cognitive apparatus more fit for what morality demands. This includes making use of one’s feelings of love in the service of moral reason, and in this general sense the ‘must be rational’ can be interpreted as involving a demand to cultivate natural feelings of love of benevolence and love of delight in order to improve oneself morally (see esp. ch. 4.2; ch. 5).

Even though Kant’s published works contain only one direct reference to love’s being ‘generally’, ‘as such’, or ‘all-in-all’ [überhaupt] divided into love of benevolence and love of delight, his usage of the terms of the division runs from the early Herder lectures on ethics (1762–64) up to the late Metaphysics of Morals (1797) (see ch. 4). The first time the general division of love comes up explicitly in Kant’s corpus is in the Collins notes on ethics, where ‘all love’ [alle Liebe] is divided into love of benevolence and love of delight (LE, 27:417.19–30).² There, the specific context is love of neighbour. With respect to neighbourly love, in the 1780s Kant also uses the distinction between pathological and practical love, which is especially familiar to readers of the Groundwork.

² Because the manuscript of these notes is dated 1784/85, it is often thought that they are roughly contemporaneous with the Groundwork of the Metaphysics of Morals. However, as the notes are nearly identical to earlier ones ascribed to Kaehler, it has been established that the Collins notes in fact stem from the mid-1770s, which explains some of their features (like their emphasis on inclinations) compared to Kant’s mature moral theory (see Denis/Sensen 2015, pp. 3–4; Naragon 2006, entry on ‘Collins 2’ in ‘Moral Philosophy’).
The division of love of neighbour into love of benevolence and love of delight resurfaces in *The Metaphysics of Morals*, and in ch. 4 I investigate how the general division of love of neighbour can be mapped onto the more familiar pathological-practical distinction. If all the evidence is taken into account, including what Kant says about love in friendship (ch. 5), we can see that love for other human beings is generally divided into love of benevolence and love of delight, so that it is possible to love others benevolently 1) from inclination or 2) from reason, and to take delight in others 3) pathologically or 4) intellectually (see ch. 4.1; 4.3; 5.2.2; Appendix). From a religious perspective, God’s love of benevolence is the ground of creation and moral duties, and his love of delight is (hopefully) an eventual favourable response to the sincere moral striving of the human being (ch. 3). How exactly the general division of love in Kant operates is to a great extent the main problem of the whole study, and my exegetical work is largely meant to corroborate the existence of the general division of love in Kant. The investigation of the general division of love reveals that the aspects of love are not isolated from each other but overlap to some extent and, taken together, form a dynamic and highly complex network of closely intertwined concepts. It should also be noted that *Kant’s conception of love cannot be entirely reduced to the general division*: under the broad and complex rubric of self-love, we find the strongest, rudimental, non-rational impulses of human nature, namely those of self-preservation (love of life) and sexuality (sexual love in the narrow sense), which are discussed in terms of love but not grasped by the general division (ch. 1.1; 2.2.1).

The second major claim of the study is that if we look at the different aspects of love alongside each other, we see an *ascent of love* from the natural, animal

---

3 The distinction between love of benevolence and love of delight has its historical roots, of course, but as I already mentioned, in this work I cannot offer detailed historical comparisons or a proper account of the genealogy of this pair of concepts prior to Kant. A similar distinction appears, for instance, in Hutcheson’s *An Inquiry into the Original of our Ideas of Beauty and Virtue*. There, Hutcheson divides love toward other rational agents into love of complacency and love of benevolence (see Hutcheson 1990 [1725], pp. 127 ff., Treatise II.2–5), which were translated into German as *Liebe aus Wohlgefallen* and *Liebe aus Wohlwollen*. Apparently, Kant owned the German translation of Hutcheson’s work (I thank Michael Walschots for providing me with this information regarding the link between Hutcheson’s and Kant’s conceptions of love). More generally, *amor benevolentiae* and *amor complacentiae* belonged to standard Scholasticism and were discussed at length by Aquinas in the *Summa Theologiae*. In his highly intricate conceptual network of love, Aquinas includes a third general notion: *amor concupiscientiae*, or love as desire. Through Aquinas, the origin of the general division of love can be traced to Christian theology and Aristotle’s concept of benevolence, especially in his discussions of *philia* in the *Nicomachean Ethics*. 
impulses toward the highest moral-physical good in the form of cosmopolitan friendship – or so I will argue. Through the aspects of self-love, sexual love (and love of beauty), love of God, love of neighbour, and love in friendship, love is seen to condition important focal points in humanity’s ascent from crude animality to morally deserved happiness. This picture of an ascent of love in Kant is obviously an interpretative reconstruction. Kant never systematised his discussions of love into a single whole, but the reconstruction I offer is nevertheless based firmly on what he said. It is made from the pieces Kant laid out, even though he himself never put all the pieces together. The picture contains both descriptive elements, which Kant uses to portray human nature, and prescriptive elements, which explicate notions of duty as they relate to love. The ascent concerns both the subjective level of an agent’s character development and the communal level of the species. It also contains the regulative ideal of the highest good as the perfection that we humans ought to strive for. Overall, this view of an ascent of love is a conceptual classification or a hierarchy of the different kinds of love as they relate to creation, nature, and the highest moral-physical well-being. I call this picture the ascent model of love in Kant. The model consists generally of the various notions of love spelled out by Kant in his works, the interrelations between these notions, and the way in which the general division of love brings a relative unity to Kant’s concept of love as a whole. The ascent model of love ultimately provides a panoptic view of the aspects of love discussed in this book. To make the claim more precisely, the ascent model of love is a viable general model of love in Kant.

For readers familiar with Plato’s Symposium and the famous ‘ladder of love’ discussed in that work, the notion of an ascent model might ring a bell. Isn’t Diotima’s and Socrates’ account of Eros in the latter half of the Symposium precisely an ascent model? Am I trying to argue that Kant is actually some kind of Platonist when it comes to love? The answer to the first question is: yes, the first ascent model of love in Western philosophy was formulated by Plato. To the second question, I’m inclined to answer no, but this must be carefully qualified to avoid misunderstandings. Naturally, if I am generally arguing that an ascent of love can be detected in Kant, there are going to be at least some structural similarities with Plato’s (or Diotima’s) account in the Symposium. With Plato, one begins by erotically loving the physical beauty of an individual young man, and the impulses of self-preservation and the sexual instinct likewise lie at the natural basis of Kant’s conception of love. Echoing Plato’s ‘ladder metaphor’, Kant often associates love with the notions of ‘higher’ and ‘lower’, and he continually talks about love in relation to our striving for perfection, or in relation to the cultivation of our faculties – where the ultimate end is obviously the complete highest good. What is clearly different, however, is that for Plato, the ‘peak’ of love’s
ascent is a kind of quasi-mystical vision, where the lover suddenly grasps the fundamental oneness of the idea of beauty, as if in a single, sweeping intuition. Now this is not a comparative study of love and the highest good in Plato and Kant; it is not even a study of the highest good in Kant, and I will not go into much detail in my comparison. Using Plato’s ladder casually as a heuristic point of departure through which to elaborate on Kant’s notion of love, I believe we can say that for Kant, the ‘peak’ we strive to reach is a more communal notion of love (conditioned by respect) in cosmopolitan friendship. Subjectively, it consists in *love for the moral law*, the full attainment of which signifies the absence of all contra-moral inclinations in the agent. Communally, or on the species level, it consists in the prevalence of benevolence (*Liebe des Wohlwollens*) and intellectual delight (*Liebe des Wohlgefallens*) in equal, reciprocal, and respectful human relationships that ultimately obtain throughout the planet. For Kant, the attainment of these highest modes of love is a gradual, laborious, and never-ending project of moral development. The Kantian agent is never ‘rewarded’ with the actualised, sweeping vision of the beautiful, radiant oneness described by Plato. For Kant, the ascent of love is less about subjectively coming to ‘see’ something and more about making moral progress in terms of love in one’s interactions with other human beings. I am not saying that Plato’s account cannot also be construed along these lines, but on the face of it at least, there are clear differences between the two. Plato’s ladder of love emphasises the vision of the one as the highest good, whereas my ascent model of love in Kant emphasises the duty of moral progress.

**IV**

If I now briefly compare my perspective with the previous, more general propositions made on Kant and love, I believe the benefits and originality of my approach can be brought to light. First of all there is an older doctoral dissertation (55 pages) from Germany with the title *The Concept of Love According to Kant [Der Begriff der Liebe bei Kant]* (Streich 1924). Detlev Streich’s main claim is that love can never be a moral motive for Kant. While this is strictly speaking true, Streich’s coarse-grained position reduces love to a feeling, and he does not problematise his conception in the light of Section XII of the Introduction to the ‘Doctrine of Virtue’ (see Streich 1924, p. 38), where the feeling of love is described as a subjectively necessary predisposition for receptivity to duty. Streich does not discuss love of God or analyse the varieties of self-love; he only mentions sexual love in passing (Streich 1924, p. 46) and has a particularly one-sided view of love in friendship (he thinks it’s merely burdensome [lästig]).
(Streich 1924, pp. 42, 44). More recently, in an article on Kant and the biblical commandment of love, Martin Moors claims to ‘formulate a general evaluation of Kant’s philosophy of love’ (Moors 2007, p. 266). However, Moors does not exactly provide this, and instead identifies six aspects ‘with regard to Kant’s practical concept of love’ (Moors 2007, p. 266; emphasis added). Included in Moors’s account are ‘religious’, ‘theological’, ‘theonomical’, ‘ethical’, ‘voluntaristic’, and ‘anthropological’ varieties of love. From my perspective, the first three would seem to come close to love of God, the fourth and fifth to love of neighbour, and the sixth to the notion of passion. (Moors 2007, p. 266) Moors does not discuss the varieties of self-love in significant detail; he says nothing about sexual love (and love of beauty), nor does he mention love in friendship. For Moors, Kant’s notion of love as duty ‘evaporates completely’ (Moors 2007, p. 267), but he does not notice that an end of the duty of love is the happiness of others. From a more charitable perspective, Jeanine Grenberg has defended ‘a Kantian understanding of the role of love in a well-lived human life’, a love the conception of which is ‘entirely moral’ (Grenberg 2014, p. 211). Grenberg seems to effectively reduce ‘Kant’s notion of love’ (Grenberg 2014, p. 220) to love for the moral law (see Grenberg 2014, pp. 218–219; cf. ch. 3.1 below). But I think a Kantian comprehension of love’s role can hardly be reduced in this way, and besides love for the law, Grenberg provides no account of the many things Kant has to say about love.⁴ The most interesting, detailed, and modest general proposal is that provided by Schönecker (2010) in the Introduction to his paper on love in Section XII of Kant’s Introduction to the ‘Doctrine of Virtue’. First of all, Schönecker makes clear that his list is not perfect, and the aim is only to demonstrate the complexity of the concept of love (Schönecker 2010, p. 135). Schönecker divides love in Kant into at least four contexts and twelve different meanings. The first context is biological and includes 1) sexual love, 2) self-love, and 3) love of life (self-preservation). The second context is the duties of love, where Schönecker identifies 4) amor benevolentiae, 5) heartfelt benevolence, 6) love for all human beings, 7) love as an aptitude to the inclination of beneficence, and 8) practical love. Schönecker’s third context is love in friendship, where he distinguishes between 9) love as friendship with humanity, 10) the duty of benevolence as a friend of human beings, and 11) benevolence in wishes. The last context is love as a moral predisposition of amor complacentiae, which is 12) love of delight [Liebe des Wohlgefallens]. Without going into de-

⁴ Perhaps it would be more charitable to view Grenberg’s account as a work in progress, because elsewhere she also discusses love in the Herder and Collins notes on ethics (see Grenberg 2015).
tail, Schönecker’s analysis is very helpful, and from my perspective the only main contexts that are omitted from his account are love of God and love of beauty. As Schönecker does not ground his tentative analysis on the general division of love into love of benevolence and love of delight, it is understandable that he is not so sensitive to the operation of the division within the contexts he distinguishes. I think Schönecker’s loves 4 – 11 can actually be viewed in terms of varying kinds or degrees of love of benevolence. Love of delight, on the other hand, is a broader notion than Schönecker acknowledges in his list; it also figures in self-love, love of God, and, arguably, in love in friendship. The context of biological or natural love can also be construed such that the non-rational impulses of love of life and sexual love (the latter of which, in the broad sense, involves more than just biology) belong to an umbrella concept of self-love, which further includes the general division of love on an actually rational level (see ch. 1).

V

This study is divided into 5 chapters according to the aspects of love on which I will focus. In these chapters, I analyse the operation of the general division of love in the relevant contexts and carve out the various building blocks that, taken together, form the ascent model of love in Kant. Kant’s discussions of love take place mainly within moral philosophy, philosophy of religion, anthropology, and teleology, while he does not really talk about love in the framework of theoretical philosophy (‘love’ is not mentioned in the Critique of Pure Reason). Apart from the first Critique, my readings emphasise the main published works from the mature period, and I only use the lecture notes and minor writings as auxiliary tools of interpretation when helpful. My reading strategy is fairly consistent from chapter to chapter; while I tend to organise my discussions thematically around the problems related to the general division of love, I normally read Kant’s works chronologically to appreciate the transformations his thoughts on love undergo over time. I believe that, like any great system or project of human thought, Kant’s philosophy must be understood as a dynamic and cumulative endeavour, where more stable positions are found on which details are then built, while some positions change, wane, or become redefined, and the thought itself remains constantly at work, constantly in flux. Following the methodological advice of Ernst Cassirer, mine is not so much a study of ‘puzzles’ or ‘apparent contradictions’ in the 29 volumes of Kant’s collected works but rather a study of the dynamic structure of a particular philosophical concept – a concept which a philosopher can only form gradually over the course of several decades
of conceptual labour (see Cassirer 1981, pp. 1–2, cf. 137–138). Such a concept of love, while certainly not the only possible such concept, can nevertheless be very beautiful, and can perhaps help us more generally to understand what love is, how it arises, what it feels like, and what it requires from us rationally. To understand love better is the deeper, underlying aim that has resulted in this book.

In the first chapter, I offer a three-level interpretation of self-love in Kant. I argue that the concept of self-love can be divided according to ascending levels of rationality. I show that there is a low, arational, ground-level form of self-love that consists of the strongest animal impulses of human nature: love of life (self-preservation), sexual love and parenting (species preservation in the narrow sense), and instinctive sociality. I call this level ‘animal mechanical self-love’.

My analysis then turns to an actually rational level that I call the ‘middle level’ of self-love, where self-love is divided (in a very complex way) according to the general division of love into love of benevolence and love of delight. Lastly, I analyse self-love hypothetically, on an ideally rational level of infinite approximation towards the highest good. I argue that even in this infinite ascent an element of self-love persists. Overall, I thus argue for the persistence of self-love on three levels.

The second chapter has two exegetical tasks. I begin by analysing the relationship between sexual love and love of beauty, which relationship figures prominently in Kant’s earlier philosophy but wanes towards the 1790s. I then formulate a distinction between narrow and broad sexual love, where narrow sexual love consists merely of the natural impulse of procreation, whereas broad sexual love is the natural impulse united with the moral love of benevolence in the context of heterosexual marriage. In comparison with the other chapters in this study, the chapter on sexual love also contains an evaluative element, which reflects the fact that most of the previous research in this area is feminist and therefore evaluative in orientation (as far as I know, my study is the first to analyse sexual love in Kant from an exegetical point of view). From this perspective, I show how broad sexual love supports a less misogynistic picture of Kant than is often presented, even though problems and internal tensions remain in Kant’s discussions on the issue of sex.

The third chapter formulates an ascent model of love of God. I begin with the observation that love of God comprises two ‘directions’: a movement upwards, from human beings to God, and a movement downwards, from God to human beings. I call this starting point the two-directionality thesis of love of God. I proceed by analysing human beings’ love for God and then God’s love for human beings. For Kant, morality leads to religion, and love for God is the foundation of all inner religion. It is close to the regulative ideal of loving the moral law, which involves fulfilling one’s duties gladly [gern] (thus implying the absence
of contra-moral inclinations in the perfect agent). I show how God’s love can be analysed in terms of its role as both an end and a ground: God’s love of benevolence toward humans is (from a religious perspective) the ground of creation and duties, and God’s love of delight is a moral delight God in the end hopefully takes in the sincere moral striving of the human being.

The fourth chapter proposes a novel ‘feeling-action-cultivation’ account of love of neighbour. I track down Kant’s discussions of neighbourly love from the early Herder notes onwards, focusing on The Metaphysics of Morals. I propose that love of neighbour consists of both moral-rational and sensory-emotive elements and that it includes the cultivation of a moral disposition. Love of neighbour divides into love of benevolence and love of delight, so that love of benevolence towards others is either 1) benevolent or beneficent inclination or 2) active rational benevolence (practical love). Practical love is further divided into beneficence, gratitude, and sympathy. Love of delight is either a pathological or an intellectual delight taken in the perfections (or even the sheer humanity) of another. It is not merely an actual feeling but also a predisposition of sensibility to be subjectively receptive to duty.

The final chapter reconstructs Kant’s mature philosophy of friendship from the perspective of love. I show the existence of the general division of love in this context and analyse the way its components function. Love in friendship is at least love of benevolence, but if the lecture notes on ethics are included as evidence, it is both love of benevolence and love of delight. In general I argue that in the context of friendship, love (conditioned by respect) marks the path towards the highest good in equal and reciprocal human relationships. Friendships as such are intimate, but the notion of a ‘friend of human beings’ [Menschenfreund; Freund der Menschen] brings with it a broader cosmopolitan outlook that indirectly aims at the ideal moral community in terms of friendship. I call this overall account the ‘ascent view of love in Kantian friendship’. Taken together with the previous chapters of this study, this account corroborates the general division approach and the ascent model of love in Kant.

VI

Before moving to the main discussion, I would like to say a word or two about certain aspects of love to which I do not devote entire chapters but that also deserve mention. First, I approach the notion of love of beauty through the lens of sexual love (ch. 2.1). This is by no means the only strategy available, and love of beauty could also be considered for its own sake (even though there are only two direct passages on it in the Critique of Judgment). One could, for instance, begin
one’s discussion with the third Critique, establish that love of beauty is love of
delight, and then try to establish links between these passages and Kant’s dis-
cussions in the ‘Doctrine of Virtue’ and the Vigilantius notes on ethics with an
eye to elaborating further on the general qualities of love of delight. But this
kind of more experimental exegesis is beyond the scope of the present work.
Love of beauty has been recently discussed by Anne Margaret Baxley (2005)
and Gabriele Tomasi (2015), and with reference to their discussions I adopt a
merely reactive attitude in ch. 2.1, pointing out that an exegetical problem con-
cerning love in the Critique of Judgment is left untouched by these accounts.
However, this is not the main point of the chapter, and in all honesty I do not
provide a general account of love of beauty.

Second, I discuss love of honour under self-love. Love of honour is an ambig-
uous notion detached from the general division of love (in itself it is neither love
of benevolence nor love of delight). Love of honour consistently marks a concern
for respect (from others), but it comes in physical and moral variants, the phys-
ical variant belonging to the conceptual cluster of self-love and the moral variant
being grounded in respect for the moral law. My discussion of love of honour is
indebted to the careful accounts provided by Houston Smit and Mark Timmons
(2015) and Lara Denis (2014), but I should still note that neither Smit and Tim-
mons nor Denis attempt to connect love of honour to love’s general concept. I
consider the relationship between love of honour and love in general at the
end of ch. 1.1. The moral ideal of love for the law is discussed in ch. 1.3 and es-
pecially in ch. 3.1. I consider love of human beings in conjunction with love of
neighbour, especially in ch. 4.2.1.

Further, there are various ‘aspects’ of love that are mentioned by Kant only
in passing, in singular isolated contexts or in adverbial constructions, but that
are never elaborated on or systematically developed. Some of these might be
more, some less important for someone interested in a general concept of
love, but since they are not given substantial consideration by Kant in terms
of love, I do not discuss them or incorporate them into my framework. Of
such loves, the most prominent is undoubtedly parental love, yet even though
the natural impulse toward the preservation of offspring belongs to ‘animal me-
chanical self-love’ in the Religion (R, 6:26.12–18), the only published reference to
‘parental love’ [die Liebe der Eltern] occurs in the Prolegomena, where Kant uses
love of God and parental love as examples through which to explain, formally,
the notion of an analogical relation as such (4:357fn.; cf. 28:1087–1088). Parents
do have a duty to provide for their children, according to Kant, and children are
said to have a duty of gratitude (which is a duty of love) towards their parents
(MM, 6:280–281; cf. 9:482). When Kant discusses the difference between hatred
and anger in the lectures on ethics, he mentions a parent’s anger toward a
child’s bad behaviour as an example of anger that presupposes love (LE, 27:687.38–688.1). That parents love their children would seem to be implicit in Kant’s writings, but he does not discuss the parent-child relation in connection with love to any great extent. Hence, I do not include parental love in my framework.

Love of truth is mentioned a couple of times in the lectures but never in the published works (see LE, 27:60–62; 27:448–449). In the first Critique, Kant writes that ‘we shall always return to metaphysics as we would to a beloved woman with whom we have had a quarrel.’ (C1, A850/B878)⁵ Although this is an interesting metaphor, to my knowledge Kant never elaborates on it. Even more remote examples of ‘briefly mentioned loves’, which for the most part never appear in Kant’s published works, include ‘love of the fatherland’, as contrasted with universal love of human beings (LE, 27:673; see my ch. 5.3), ‘love of justice’ (LE, 27:688–689), ‘peace-loving’ (C2, 5:61.10; see also LE, 27:687.2), and the carnivorous ‘love of roast beef’ (LA, 25:1361.8), none of which are developed further in terms of love. The existence of constructions like ‘love of roast beef’ merely shows that, in the most general terms, ‘love’ can be used to signify any kind of relatively intense liking or desire. From this flexibility of the concept of love it does not follow that the framework of love should be expanded ad infinitum to accommodate ever-new aspects or kinds of love. Rather, it shows the need to restrict the framework through careful, quasi-inductive evaluations, so that the concept of love can remain at the same time broad and informative.

⁵ ‘Man kann also sicher sein, [...] man werde jederzeit zu ihr wie zu einer mit uns entzweiten Geliebten zurückkehren’.
1 Self-Love

According to Kant, it is a basic fact of human nature that we all love ourselves. Human beings are imperfect rational creatures who want things to go well for themselves, and there is really no way around this fundamental trait of our species. We want to be happy. Indeed, Kant tends to describe self-love [Selbstliebe, Eigenliebe] as the basic principle of subjective happiness. Self-love is the natural motivational ground of action, to which the moral incentive (respect for the moral law) is opposed. Self-love is active within us, and it is only hindered by morality to the extent that it gives rise to contra-moral inclinations. There are duties in Kantian ethics, of course (e.g. the duties to perfect oneself and to promote the happiness of others), and in light of our duties we should keep self-love in check and diminish its influence. Self-love can often appear as selfishness, and as such it is prone to impede our striving for virtue. It may threaten the freedom and happiness of others.

But what does the above actually mean, and can something like this be said to be the whole picture of self-love in Kant? While the basics are relatively simple from the perspective of moral theory, self-love poses problems for Kant’s ethics. It is a significant concept in his moral philosophy and has received attention from scholars. Treatments of self-love in the previous literature, however, can be quite varied. Some view self-love as ‘an objective principle of practical reason’ (Paton 1947, p. 91) or even as ‘furthering morality’ (Šimfa 2013, p. 107), whereas others hold that ‘love is not an attitude that clear-sighted and rational people could ever take toward themselves.’ (Wood 1996, p. 144) Often, scholars touch upon self-love in discussing other issues, such as beneficence (Hill 1993), benevolence (Edwards 2000), or respect for the moral law (e.g. Reath 2006; Engstrom 2010). Yet none of the previous readings aim to provide a systematic, exegetical interpretation of the notion of self-love.

1 I have been unable to detect a difference in meaning between the two German terms, even though it seems that Kant tends to use Eigenliebe when he is contrasting self-love with self-conceit [Eigendünkel] (see e.g. C2, 5:73.12–14).
2 Thomas Hill’s focus (1993, pp. 1–2) is generally on the possibility of altruism, and he draws from a loosely Kantian framework, stating explicitly that his aims are not exegetical. Jeffrey Edwards’s reading focuses on the Groundwork and the second Critique and is in fact a defence of Hutcheson against the basic framework of Kant’s mature moral philosophy. Andrews Reath and Stephen Engstrom analyse the moral incentive (or ‘spring’ [Triebfeder], as Engstrom puts it) and its relation to non-moral agency. In doing this, they also offer very helpful analyses of self-love, particularly in relation to self-conceit within the context of the second Critique (see Reath 2006, pp. 14–17, 23–25; Engstrom 2010, pp. 101ff.).
My fundamental goal is to understand the intricate structure of self-love and the role it plays in Kant’s thought. What, for instance, does Kant mean when he says that self-love is a predisposition to the good (R, 6:26)? Or when he argues that it is the source of evil (R, 6:45)? What does he mean when he holds that all material practical principles fall under self-love (C2, 5:22)? What is the status of rational self-love (C2, 5:73; cf. R, 6:45–46fn.), and how does self-love relate to the general division of love into love of benevolence and love of delight? Finally, what is the status of self-love in the infinite progress towards moral perfection—will it continue to exist or not?

In what follows, I shall present what I call a ‘three-level’ interpretation of Kantian self-love, according to which the concept of self-love is divisible into ascending levels of rationality. I begin by discussing self-love at a rudimentary, non-rational level of the cognitive structure of the human being, where it figures as the strongest impulses of human nature. These fundamental animal impulses include self-preservation (love of life), preservation of the species (sexuality and parenting), and sociality. Together, they may be identified as ‘animal mechanical self-love’, and according to Kant they constitute a predisposition to the good. The main point of the first section is to lay the ground for the three-level interpretation of self-love (and for the ascent model of love in Kant more generally) by showing that Kant discusses the strongest impulses of human nature in terms of love. I will strive to understand why Kant calls these non-rational animal impulses ‘love’ in the first place, and why he thinks that they constitute a predisposition to the good. To this end, I will also problematise the relationship between animal self-love and the ‘self-love of humanity’ (the latter of which implies what I call the ‘middle level’ of self-love) by looking at notions of sociality and love of honour in this context.

Second, I analyse the middle level of self-love in more detail, from the perspective of Kant’s moral philosophy. I argue that this level of self-love is best approached by acknowledging the operation of the general division of love as love of benevolence and love of delight. This level implies the actuality of reason, and here self-love can be considered to inform all non-moral ends the agent may have and the instrumental reasoning related to them. The middle level brings with it the notion of self-conceit, which Kant incorporates into the framework of self-love through very complicated discussions. Love of benevolence for oneself is willing one’s own happiness (or love from others), and this is a permanent and acceptable part of humanity; self-conceit, by contrast, is a morally reprehensible, delusional delight [Wohlgefallen] taken in a sense of special self-worth in comparison with others. Self-conceit arises when one makes the self-love of benevolence an unconditional law. Rational self-love, on the other hand, refers to
the middle level of self-love under moral conditions: it is prudence limited by morality (love of benevolence) or moral self-contentment (love of delight).

Third, I consider what would happen to self-love in the ideally rational state of moral existence, where the highest good as moral happiness is realised as closely as possible. Here, I argue that because the end of morality will necessarily involve happiness, it follows that self-benevolence and moral self-contentment will be present in the infinite approximation towards the highest good. I don’t see a way to conceive of morally deserved happiness without including some kind of self-benevolence or moral self-contentment, both of which can arguably be cashed out in terms of self-love. Hence, the three levels of self-love in my interpretation consist of: 1) animal mechanical self-love; 2) the middle level or general division of self-love; and 3) self-love and the highest good.

I thus argue that while there is an important sense in which self-love is in tension with morality and constitutes an obstacle to moral progress, it is an irreplaceable component of human existence. In Kant’s moral thought, there is no prescription (or even possibility) of a type of agency that is completely stripped of self-love (even though unselfish acts may very well be commanded). Self-love cuts through the ascending levels of the cognitive structure of the moral agent and can also be used to illuminate the notion of moral progress from a broader species perspective. Although the third and highest level of self-love is of course idealised, it functions in my argument to illustrate the persistence of self-love in the infinite approximation to the highest good. No matter how closely the moral happiness of all rational creatures is realised, the Kantian agent will still retain some minimum of an attitude of love towards herself.

Overall, then, the three-level interpretation of self-love has two exegetical functions: 1) it outlines the first relatively comprehensive analysis of the conceptual structure of self-love in Kant, and 2) it serves as a preliminary for the ascent view of Kant’s conception of love as a whole.

1.1 Animal Mechanical Self-Love

We cry for food and flee from fire. We crave sex, even without seeing or thinking of anyone in particular to have sex with. We will do almost anything to keep our children alive. We are drawn to others of our kind for warmth, shelter, and acceptance. These notions clearly express fundamental human desires – desires that members of our species tend to share. But are these desires expressions of love? And if so, how?

Kant does think that crude self-preservation, sexuality and care for offspring, and our instinctive attraction to other human beings can be discussed in terms of
love. But these loves, at their natural core, are not Kant’s usual concern when he talks about love. They are not his concern when he writes in the Religion: ‘Like love in general, self-love too can be divided into love of benevolence and love of delight (BENEVOLENTIAE ET COMPLACENTIAE), and both (as is self-evident) must be rational.’ (R, 6:45.22–25) As such, the crude natural loves, as fundamental impulses of desire, operate completely irrespective of reason. They are non-rational, and they are neither benevolence nor delight. Given that the first major claim of my book is that the general division of love into love of benevolence and love of delight is a key to understanding love in Kant, it may seem peculiar that I begin the overall argument by pointing to a love (or a set of loves) that is not grasped by the general division of love. I seem to begin with a counter-example to what I wish to argue for. But while the general division of love is indeed a key to understanding love in Kant, it does not follow that the concept of love in Kant is completely reducible to the general division. This leaves us with the problem of how to understand those loves that Kant discusses as love but that do not fall into the general division of love. It is this task that I will now undertake with respect to the ‘lowest’ possible level of human existence.

The main point of the first section is to show that Kant thinks of the strongest impulses of human nature in terms of love. Together, these non-rational drives may be identified as animal mechanical self-love. But to say just that is not very conducive to understanding Kant’s view. While the main point is necessary for the argument of this chapter (and for the ascent model of love in Kant as a whole), it is equally important to ask what Kant means when he speaks of love as he does. In particular, there are two questions I find pressing in the context of animal mechanical self-love. Why does Kant call these non-rational drives ‘love’ in the first place? And what does he mean when he says that animal mechanical self-love is a ‘predisposition to the good’? In this section I provide some answers to these questions.

In the first part of Religion within the Bounds of mere Reason, Kant asserts that there is an ‘original predisposition to good in human nature’ (R, 6:26.2–3; see also 6:43.18–21). The original predisposition to the good consists of three aspects of the human being (or the whole species) in an ascending order: 1) animality (life); 2) humanity (rational life); and 3) personality (rational responsible life). The third predisposition is about morality, which is not my con-

3 ‘Wie Liebe überhaupt, so kann auch Selbstliebe in die des Wohlwollens und des Wohlgefallens (BENEVOLENTIAE ET COMPLACENTIAE) eingetheilt werden, und beide müssen (wie sich von selbst versteht) vernünftig sein.’
4 ‘ursprünglichen Anlage zum Guten in der menschlichen Natur.’
cern at this particular point, but the first two are explicitly articulated in terms of self-love. The first is about non-rational animal self-love, and the second is about a self-love for which reason is required. Because the latter is relevant to understanding the former, I will return to it later on in the section. With regard to the first predisposition, Kant writes:

1. The predisposition to ANIMALITY in the human being may be brought under the general title of physical and merely mechanical self-love, i.e. a love for which reason is not required. It is three-fold: first, for self-preservation; second, for the propagation of the species, through the sexual drive, and for the preservation of the offspring thereby begotten through breeding; third, for community with other human beings, i.e. the social drive. (R, 6:26.12–18)

As this passage clearly shows, Kant thinks that certain non-rational animal drives – self-preservation, sexuality and parenting, and sociality – are species of ‘self-love’. But why does Kant use the word ‘love’ at all in this context? He speaks elsewhere of ‘love of life’ [Liebe zum Leben] and ‘sexual love’ [Liebe zum Geschlecht] in the context of depicting ‘the strongest impulses of nature’ [die stärksten Antriebe der Natur], linking love of life to self-preservation and sexual love to the preservation of the species (AP, 7:276.28–33). In one sense it is obvious why Kant might use the word ‘love’ in the context of sexuality: there has been a close association between love and sexuality ever since the ancients, as Eros, which is traditionally used to depict sexual desire, is precisely an elementary notion of love. Another well-documented point in the literature concerning the original predisposition in the Religion is the observation that Kant owes much of his discussion to Rousseau (e.g. DiCenso 2012, p. 48; Pasternack 2014, p. 94; see also Wood 2009, pp. 127–128). In the Discourse on the Arts and Sciences, for example, Rousseau famously holds that the human being is natu-
rally good, and that developments in the arts and sciences cause moral degeneration (Rousseau 1997 [1750], pp. 4–28). In the Discourse on the Origin and Foundations of Inequality Among Men he states: ‘Man’s first sentiment was that of his existence, his first care that for his preservation.’ (Rousseau 1997 [1755], p. 161) Rousseau distinguishes between two types of self-love: _amour propre_ and _amour de soi-même_. Amour de soi-même is ‘a natural sentiment which inclines every animal to attend to its self-preservation, and which, guided in man by reason and modified by pity, produces humanity and virtue.’ (Rousseau 1997, p. 218) _Amour propre_, on the other hand, originates in society, makes people prefer themselves to others, inspires evil, and is also ‘the genuine source of honour’ (Rousseau 1997, p. 218). So there is a relatively easy answer to the question: ‘Why does Kant call these drives (self-)love, if he does not think of them in terms of his general division of love?’ The easy answer is that sexuality has been called love ever since the dawn of our culture, and the rest he just picks up from Rousseau.

But there is a more interesting answer, I think, which does not avoid making reference to the ancients but is not similarly reliant on Rousseau’s influence. The fact that Kant calls the animal impulses love may also be explained in terms of another, more general paradigm of love that has been with us at least since Plato. This paradigm of love can be called _love as desire_, and its origins trace back to Socrates’ speech in the Symposium, where the love object is by definition something that the lover desires (and, according to Socrates, lacks) (see Plato 2006, 200a–201a). For Kant, the drives of animal mechanical self-love are directly connected to what he calls the faculty of desire [ _Begehrensvormöggen_]. It is through the notion of desire and its connection to the classical paradigm of love as desire that we can shed light on the status of these impulses as love.

Further down in the Religion, Kant explains that ‘there is no question here of other predispositions except those that relate immediately to the faculty of desire’ (R, 6:28.22–23). By ‘the faculty of desire’, Kant is referring to ‘a being’s _faculty to be by means of its representations the cause of the reality of the objects of these representations._’ (C2, 5:9fn.; see also MM, 6:211; 20:206) Kant divides desire into lower and higher faculties; where the lower faculty is sensuous, the higher faculty is associated with pure practical reason (C2, 5:24–25). All living beings act in accordance with the laws of the faculty of desire (C2, 5:9fn.), and as such, animal desire is necessitated by mechanisms of nature in contrast

9 ‘hier von keinen andern Anlagen die Rede ist, als denen, die sich unmittelbar auf das Begehrensvormöggen [...] beziehen.’

10 ‘das Vermögen desselben [eines Wesens], durch seine Vorstellungen Ursache von der Wirklichkeit der Gegenstände dieser Vorstellungen zu sein.’
with the moral laws of freedom (see LE, 27:344; C2, 5:95–97). Moreover, the faculty of desire is connected to pleasure and displeasure (see LE, 27:344; C2, 5:95–97; MM, 6:211–213; cf. Frierson 2014, p. 99).¹¹ In a typical case of a ‘lower’ desire, for example a desire for chocolate (or for eating chocolate), we represent (eating) chocolate as something pleasurable, and the faculty of desire motivates us to the action of eating chocolate, or in other words, it motivates us to be the cause of the reality of the pleasure of eating chocolate. But the animal impulses go cognitively even lower than that. In Anthropology from a Pragmatic Point of View, Kant speaks of the relationship between animality and the objects of desire in terms of instincts: ‘the inner necessitation of the faculty of desire to take possession of this object before one even knows it, is instinct (like the sexual instinct, or the parental instinct of the animal to protect its young, and so forth).’ (AP, 7:265.23–26)¹² It therefore seems that the primitive nature of desire in animal mechanical self-love can be understood such that acquaintance with an object is not even required for the impulse to be operative, and even when one encounters an actual object (say food or a sexually attractive other), the satisfaction of the animal impulse does not require that one be able to form a concept of the object.

In the early ‘Observations on the Feeling of the Beautiful and Sublime’,¹³ Kant speaks of a man who loves women only as ‘things that are to be enjoyed’ [genießbare Sachen] (2:208.3–5) and who is capable of enjoying gratification

---

¹¹ For Kant’s general threefold classification of the faculties of the human mind into cognition, the feeling of pleasure, and desire, see (C3, 5:178; 20:206). Basically, through cognition we encounter objects in the first place, through desire we actualise certain objects, and through pleasure we have feelings for the objects (cf. Frierson 2014, p. 99). Frierson notes: ‘Human action is caused by desire, which is caused by pleasure, which is caused by cognition.’ (Frierson 2014, p. 99) This may be the case with the lower faculty of desire, but in the case of the higher faculty of desire pleasure cannot precede but can only follow the determination of the faculty of desire (MM, 6:212.27–213.2; see also 5:24.32–40). Frierson’s focus is on empirical psychology, which may explain why he relates the higher faculty of desire not to pure practical reason (C2, 5:24.35–36) or pure rational principles (MM, 6:212.31) but more loosely to ‘character’ and ‘maxims’. According to Frierson, the higher faculty of desire can have empirical determining grounds, as in the case of a ‘person who smokes [cigarettes] from principle’ (Friersen 2014, pp. 99–100). This reading contradicts Kant’s account of the higher faculty of desire as laid out in his moral philosophy (cf. MM, 6:426.20–26).

¹² ‘die innere Nöthigung des Begehrungsvermögens zur Besitznehmung dieses Gegenstandes, ehe man ihn noch kennt, der Instinct (wie der Begattungstrieb, oder der Älterntrieb des Thiers seine Junge zu schützen, u. d. g.).’ Note that here, the ‘object’ seems to refer to another living creature as such rather than the pleasure taken in interacting with it. This is an ambiguity that Kant does not clarify in his discussions of desire.

¹³ ‘Beobachtungen über das Gefühl des Schönen und Erhabenen’.
'without ever having to envy others or even being able to form any concept of others’ (2:208.6–9)\(^4\). The feeling this man enjoys ‘can occur in complete thoughtlessness’ (2:208.14–15)\(^5\). Despite the slight difference in terminology, the example seems to match what Kant says about instincts in the *Anthropology* quite closely. As non-rational and instinctive, animal mechanical self-love (as in the case of the sexual impulse) operates thoughtlessly, without acknowledging the personhood of others, indeed even without a concept of others. In terms borrowed from Rae Langton, animal mechanical self-love can give rise to *solipsism* (Langton 2009, p. 316, see also pp. 325 ff.)\(^6\), where the agent lets the rudimentary desires intrinsic to animal life override reason and understanding and ends up taking no account of the ends or the personhood of others, or her own personhood. This is how various vices can become grafted onto animal mechanical self-love.\(^7\)

The animal within us is consumed by desire. It is drawn to others, but it does not understand that others exist as separate from itself, and it knows nothing of the moral vocation of the organism of which it is a constitutive part. The animal within us does not understand that others set ends autonomously and that they have a capacity to be happy. This helps to explain what puts the *love* in animal

---

\(^4\) ‘ohne daß sie andere beneiden dürfen oder auch von andern sich einen Begriff machen können’.

\(^5\) ‘bei völliger Gedankenlosigkeit statt finden können.’

\(^6\) Langton uses the term mainly in the context of sexuality and attaches it specifically to some of Kant’s discussions of sexual desire or sexual love (see *LE*, 27:384 ff.; Langton 2009, p. 316). I discuss the issue of sexuality further in ch. 2.

\(^7\) These are vices of crudeness [*Rohigkeit*], such as gluttony [*Völlerei*], lust [*Wollust*], and wild lawlessness [*wilde Gesetzlosigkeit*] (see R, 6:26.18–27.3; cf. MM, 6:424–427). Because I have chosen to outline the general conceptual structure of self-love, and because my main focus is understanding the broadest conceptual divisions of love related to it, it is not possible for me to analyse specific vices (or the duties to which they are opposed) in detail. Since ‘self-love’ is the general term in Kant’s moral philosophy to which the moral incentive and the overall framework of duties is contrasted, the points of contact between self-love and duties would be overabundant considering my aims. There are specific negative duties against suicide, unnatural sex, gluttony, and the misuse of substances, which are connected to the animal nature of the human being (MM, 6:421–427). Kant insists that these duties are grounded in the categorical imperative (not animal nature) (see MM, 6:422.31–423.6; 6:425.23–26; 6:427.5–19; *GW*, 4:429.15–25; 4:425.12–27; 4:421.24–423.14). The success of his various arguments has been widely debated in the literature (for suicide, see e.g. Paton 1947, pp. 150–154; Korsgaard 1996, pp. 87–92; Wood 1999, pp. 84–86; Allison 2011, pp. 183–184; cf. Timmermann 2007, p. 81; for unnatural sex, see Denis 1999 and Soble 2003). Here, I cannot add anything new to these debates, and a proper attempt to do so would throw me off course with respect to my chosen focus on seeking out and clarifying the most general divisions of (self-)love in Kant.
mechanical self-love: animal mechanical self-love is a form of self-centred or species-centred desire that does not take others into account. It is a kind of love but not a case of ‘true love’ [wahre Liebe] (see 8:337.33–34), which would require reason and respect.

We know that Kant grounded the moral good in pure practical reason and viewed rationality as that which distinguishes humanity from animality. How, then, is animal mechanical self-love a predisposition to the good? Scholarly discussion on this topic would appear to be scarce. In particular, previous commentators do not problematise the notion of a predisposition to the good in relation to animal mechanical self-love; they simply tend to assume that animal mechanical self-love is good without explaining how this could be so.¹ Before moving on to analyse the middle level of self-love in more detail, let us pause to consider this question.

‘Predisposition’ [Anlage] is originally a term from biology (see Shell 2015, p. 96). For Kant, a predisposition is a natural feature, property, or capacity of an organism (or species) that accounts for its developing in a certain way (see Allison 2009, p. 26fn.; cf. Wood 1999, p. 211; 2009, p. 113; see also 2:434–435; cf. MM, 6:399.11). In the Religion, Kant writes: ‘By the predispositions of a being we understand the constituent parts required for it as well as the forms of their combination that make for such a being.’ (R, 6:28.17–19)¹⁹ If the being is not even possible without a given predisposition, then the predisposition is original, and if this is not the case then the predisposition is contingent (ibid., 6:28.19–21). As original predispositions, animal self-love, the self-love of humanity, and personality are all fundamental to the human species. As already noted, even though the predispositions are regularities, they are not static. They are not ready-made, rigid ‘building blocks’ of organisms but instead involve developmental processes. According to Kant, it is a general feature of nature (including human history) that ‘[a]ll natural predispositions of a creature are determined

¹ For instance, Gordon Michalson claims that: ‘All three predispositions [...] are good in themselves’ and that they ‘effectively constitute the hand we are initially dealt, while [...] how we play it depends upon the way we ourselves introduce the wild card of freedom.’ (Michalson 1990, pp. 39–40) Stephen Palmquist writes: ‘Animal self-love predisposes living beings to do good by causing them to preserve themselves, propagate the species, and form social groups for mutual protection.’ (Palmquist 2009, p. xxiv) In the same vein as Michalson, James DiCenso holds that Kant’s ‘discussion [of the predispositions] starts from an assumption of the goodness of our nature as such’ (DiCenso 2012, p. 47), whereas Lawrence Pasternack notes merely that ‘the Pre-disposition to Animality has a kind of innocence’ (Pasternack 2014, p. 94).

¹⁹ ‘Unter Anlagen eines Wesens verstehen wir sowohl die Bestandstücke, die dazu erforderlich sind, als auch die Formen ihrer Verbindung, um ein solches Wesen zu sein.’
sometime to develop themselves completely and purposively.’ (8:18.19–20)²⁰ In the human species, the unfolding of our higher, rational predispositions can only take place through a gradual, intergenerational, and indefinitely long cultural process (8:18–19). As Allen Wood and Paul Guyer remind us, Kant conceives of the natural predispositions of the human species teleologically (see Wood 1999, pp. 209–211; Guyer 2009, p. 145). The mature Kant views nature as if it were a system of ends or purposes, where the natural predispositions of the human being are ultimately referred to our moral vocation, which is connected to the ideal of the complete highest good (the morally deserved happiness of rational creatures) as the final purpose of the world (see e.g. C3, 5:429.29–32; 5:434–435; 5:451). Naturally, the highest good of the human species cannot be reached in a single lifetime or by an individual organism. At best, the species can gradually approximate the highest good.

But even if we reflect on animal mechanical self-love from the perspective of our moral vocation, its goodness, or the goodness to which it is predisposed, may still appear problematic. What exactly is the function of this self-love with respect to the good? In the second Critique, Kant distinguishes between two notions of the ‘good’ in the German language, using ‘das Wohl’ for physical well-being or pleasure and ‘das Gute’ for the moral good (C2, 5:59–60; see also 5:62–63).²¹ In the Religion (written after the second Critique), it is ‘das Gute’ that appears in the heading of Kant’s discussion of the original predisposition. But as animal mechanical self-love is physical and not moral, in the light of the second Critique it should be impossible for animal mechanical self-love to be good [gut] ‘in itself’ or ‘as such’, against the claims of some commentators (see Michalson 1990, p. 39; DiCenso 2012, p. 47). Indeed, Kant is not making the unqualified claim that animal mechanical self-love is good [gut] as such. Here is Kant’s own clarification: ‘All these predispositions in the human being are not only negatively good (they do not conflict with the moral law) but they are also predispositions to the good (they further compliance with that law).’ (R, 6:28.12–14)²² ‘The good’ of which Kant speaks here is clearly the moral good, but it is not the case that animal mechanical self-love is (morally) good in itself. Kant does not elab-

²⁰ ‘Alle Naturanlagen eines Geschöpfes sind bestimmt, sich einmal vollständig und zweckmäßig auszuwickeln.’
²¹ Note, however, that even here, where he explicitly defines the concept of the good [das Gute] in terms of morality, he allows the use of ‘good’ [gut] as an adjective in the context of instrumentally rational action from empirical determining grounds (C2, 5:62.30–31).
²² ‘Alle diese Anlagen im Menschen sind nicht allein (negativ) gut (sie widerstreiten nicht dem moralischen Gesetze), sondern sind auch Anlagen zum Guten (sie befördern die Befolgung desselben).’ Translation modified following Pluhar.
orate on what he means by ‘furthering’ [befördern] the moral good in terms of animality, but the idea must be that animal mechanical self-love somehow helps, promotes, or facilitates moral progress or the formation of our moral character on the species level. But how?

Naturally, animal mechanical self-love is a necessary condition for any kind of progress on the part of the human species – it concerns the very possibility of our species as such. If the animal within us didn’t struggle for its own preservation, if it were repulsed by all representations of sex, if it couldn’t care less about its offspring (if any even appeared), and if it were prone to isolate itself completely from others of its kind, there would be no future for the human being. In this fundamental and necessary sense, animal mechanical self-love enables moral progress.

Admittedly, if the published Anthropology is included as evidence, Kant’s use of the notion of the good in conjunction with animal mechanical self-love is more ambiguous than might be implied by the Religion. In contrast to the second Critique’s distinction between the physical ‘das Wohl’ and the moral ‘das Gute’ (C2, 5:59–60), the published Anthropology divides ‘das Gute’ into physical and moral goodness (AP, 7:277–78)²³. There, in a section titled ‘On the Highest Physical Good’ [von dem höchsten physischen Gut], the animal mechanical impulses of love of life (self-preservation) and sexual love (species preservation) are connected with what is physically best for the world [das physische Weltbeste] (AP, 7:276–277.3):

The strongest impulses of nature are love of life and sexual love, which represent the invisible reason (of the ruler of the world) that provides generally for the highest physical good of the human race by means of a power higher than human reason, without human reason having to work toward it. Love of life is to maintain the individual; sexual love, the species. (AP, 7:276.28–33)²⁴

---

²³ Since the Anthropology contains material from several decades of Kant’s thinking, one could conjecture that this distinction is a remnant of an earlier period of Kant’s thought. Yet the sections that contain this use of ‘das Gute’ in the published works are not to be found in Kant’s earlier lectures on anthropology. This suggests that the material was added, or at least architectonically reorganised, well after the publication of the second Critique.

²⁴ ‘Die stärksten Antriebe der Natur, welche die Stelle der unsichtbar das menschliche Geschlecht durch eine höhere, das physische Weltbeste allgemein besorgende Vernunft (des Weltregiers) vertreten, ohne daß menschliche Vernunft dazu hinwirken darf, sind Liebe zum Leben und Liebe zum Geschlecht; die erstere um das Individuum, die zweite um die Species zu erhalten’. Note that it is somewhat unclear whether unsichtbar refers to the invisibility of the reason or to the invisibility of the process the reason generates.
This does not have to be viewed as conflicting with the account in the *Religion*, but it adds another dimension to how animal impulses are predisposed to the good. The physically good must be understood in conjunction with pleasure. In the light of this evidence there is also a sense in which the strongest impulses of human nature (love of life and sexual love) are good because they give us pleasure, apparently via good meals (see AP, 7:278.10–12) and sexual intercourse.²⁵

As already mentioned, no progress whatsoever is possible on the species level without animal mechanical self-love. This is the bare minimum that Kant must have in mind with respect to the relationship between animal mechanical self-love and the moral good.²⁶ There may, however, be stronger connections. I discuss the issue of sexual love further in ch. 2, but there is another link that can shed light on animal self-love as predisposed to the good. This link is the animal impulse of sociality [*Gemeinschaft; Gesellschaft*] (R, 6:26.17–18)²⁷. The so-

---

²⁵ This passage may also raise questions about the role of God’s reason in this context. As we are here dealing with a physical good (pleasure) that human reason does not have to work for (AP, 7:276.30–31), it seems that Kant’s reference to God’s reason shouldn’t (at least on the face of it) be primarily understood from the perspective of the moral postulate of God. Rather, the reference seems to point toward the notion of providence [Vorzuschung], which Kant makes use of in the philosophy of history (see e.g. 8:30.19; 8:121.3; 8:123.23). Interpreted from this perspective, animal mechanical self-love yields physical pleasure according to a ‘hidden plan of nature’ [*der verborgene Plan der Natur*] (see 8:272–273), the final purpose of which is the complete development of all our predispositions (see 8:18.19–20). Relatedly, in *The Metaphysics of Morals*, love of life (self-preservation) and sexual love (species-preservation) are connected with ‘an intelligent cause’ [*der Ursache [...] Verstand*] in terms of the critical teleology (see 6:424.12–18). In the critical teleological framework articulated in the third *Critique*, all natural purposes must be thought of as if they were intelligently caused (even though we cannot know this theoretically) if they are to be systematically connected to the ideal or the final purpose of the highest good (the morally deserved happiness of rational creatures). The reason for this is that Kant thinks that the highest good is possible only if God is postulated (see ch. 3).

²⁶ In Kant’s *Lectures on Pedagogy*, the first principle of education is to discipline the wildness of animality, and hence the education related to animality is negative: ‘To discipline means to seek to prevent animality from doing damage to humanity, both in the individual and in society. Discipline is therefore merely the taming of savagery.’ (9:449.28–30; see also 9:441.18; 9:465.30) / ‘Disciplinieren heißt suchen zu verhüten, daß die Thierheit nicht der Menschheit in dem einzelnen sowohl als gesellschaftlichen Menschen zum Schaden gereiche. Disciplin ist also blos Bezähmung der Wildheit.’ This supports the view that animality on its own does not play a positive role in moral progress.

²⁷ In English, this animal impulse might also be termed ‘sociability’, since ‘sociality’ and ‘sociability’ are nearly indistinguishable. According to the Merriam-Webster dictionary, for instance, ‘sociality’ is ‘sociability’, but ‘sociality’ involves in particular ‘the tendency to associate
cial impulse connects animal self-love with the self-love of humanity, and looking more closely at this connection might help to explain how animal mechanical self-love is predisposed to the good. Let us therefore turn to the predisposition of humanity:

2. The predispositions to HUMANITY can be brought under the general title of a self-love which is physical and yet involves comparison (for which reason is required); that is, only in comparison with others does one judge oneself happy or unhappy. Out of this self-love originates the inclination to gain worth in the opinion of others, originally of course, merely equal worth: not allowing anyone superiority over oneself, bound up with the constant anxiety that others might be striving for ascendancy; but from this arises gradually an unjust desire to acquire superiority for oneself over others. (R, 6:27.4–12)

Kant’s discussion implies what I call the actually rational ‘middle level’ of self-love, which I discuss from a moral theoretical perspective in the second section. Here, the passage in question connects particularly well with Kant’s philosophy of history, where he uses these kinds of terms to discuss the development of human societies. Animal self-love makes us approach other animals of our kind for sex and nurturing and to live in their company, but our rational predisposition brings with it a comparative and competitive drive. As rational creatures who live in societies, we observe others and compare ourselves with them. We constantly evaluate their well-being, their various skills, their wealth, power, how good they are morally, and so on. We do not want to be worse than others, and we are afraid that they will strive to be superior to us. We also want them to think well of us. This rational but non-moral predisposition of self-love motivates us to emulate and compete. It gives rise to social antagonism, which Kant famously calls ‘unsociable sociability’ [die ungesellige Geselligkeit] (8:20.30). In or form social groups’. Since this belongs to the animal level, I use ‘sociality’ for the animal impulse and (unsocial) ‘sociability’ for the social self-love for which reason is required.

28 ‘2. Die Anlagen für die MENSCHHEIT können auf den allgemeinen Titel der zwar physischen, aber doch vergleichenden Selbstliebe (wozu Vernunft erfordert wird) gebracht werden: sich nämlich nur in Vergleichung mit andern als glücklich oder unglücklich zu beurtheilen. Von ihr rührt die Neigung her, sich in der Meinung Anderer einen Werth zu verschaffen; und zwar ursprünglich bloß den der Gleichheit: keinem über sich Überlegenheit zu verstatten, mit einer beständigen Besorgniss verbunden, daß andere damach streben möchten; woraus nachgerade eine ungerechte Begierde entspringt, sie sich über Andere zu erwerben.’

29 The key text here is ‘Idea for a Universal History with a Cosmopolitan Aim’. Kant’s basic idea of unsocial sociability is not his own invention, and similar notions had been previously expressed by various authors. As noted above, Rousseau called the self-love related to social antagonism amour propre (see Rousseau 1997 [1755], p. 218; see also 1974 [1762], pp. 173–5 (begin-
unsociable sociability, self-love draws us together to form communities, and the same self-love also threatens to destroy them (see 8:20.30–33). Self-love can be a great source of evil, a source of arrogant self-conceit, greed, dominance, tyranny, and war, but competitive antagonism also drives culture forward and propels the development of natural human predispositions. The self-love of humanity makes us advance our various talents, taste, and the argumentative skills we need in our quest for knowledge (see Cohen 2014, p. 81). It motivates us to emulate the morality of others, and it is generally a source of our inclination for honour (see Denis 2014, p. 200). As Kant explains in ‘Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht’³⁰, this cultural process grounds a frame of mind ‘which can with time transform the rude natural predisposition to make moral distinctions into determinate practical principles and hence transform a pathologically compelled agreement to form a society finally into a moral whole.’ (8:21.14–17)³¹

This is at base how the self-love of humanity is predisposed to the good. What is particularly interesting from the perspective of animal mechanical self-love, and self-love and love more generally, is that the higher (though still physical and not moral) self-love is rooted in animal sociality. As Wood (2009, p. 115) notes, unsociable sociability is a ‘modification’ of the social animal predisposition. Being instinctively drawn to others grounds rational comparison with others, and the development of comparative self-love from this basis can somehow be morally transformative at the level of society.

Before concluding this section, let us turn to whether this idea can be made sense of by looking at the connection between animal sociality and what Kant calls love of honour [Ehrliebe]. According to Kant, the inclination to be equal to others gives rise to emulation, which we ought to cultivate and which ‘serves merely to educate our animal nature and make it adequate to humanity, or the intellectual being within us, and to its laws.’ (LE, 27:695.17–20)³² This emulative tendency gives rise to love of honour (LE, 27:695.25–30; see Denis 2014, p. 200). Drawing from previous discussions on this topic (Smit & Timmons 2015; Denis 2014), I would suggest that there are two basic kinds of love of hon-
our in Kant: one physical, the other moral.³³ The first kind is emphasised in Kant’s earlier lectures on ethics³⁴, while the latter plays a greater role in his mature moral philosophy. Physical love of honour is a natural impulse [Trieb] or inclination to secure respect and a ‘favourable judgment’ [ein günstiges Urtheil] from others (LE, 27:408.4–5; 27:408.29). Like the other drives connected to self-love and unsociable sociability, love of honour has an ambiguous nature. When we take the judgments of others into account in assessing our own actions, we indirectly pave the way for morality. This love of honour is a natural and in itself unselfish (see LE, 27:410.30–31) inclination merely not to be an object of contempt (LE, 27:408.35–37). As such, it is not directed to one’s own advantage (27:408.5–7), but it can easily turn into an ambitious craving for honour, or arrogant self-conceit, such that we come to view others as inferior and think of ourselves as being entitled to their highest respect. In Kant’s mature moral philosophy, on the other hand, it is respect for the moral law that grounds love of honour. In The Metaphysics of Morals, love of honour is a virtue based on the self-assessment of our dignity in comparison with the moral law (MM, 6:420.13–30; see LE, 27:609.5–610.9; 27:695–696; 27:667.12–21). Love of honour is also a basic claim to be respected by others because of one’s fundamental standing (and equality with others) as a moral being (MM, 6:464.5–11; see Denis 2014, pp. 206–207). The moral love of honour is not based on external concern for the opinions of others: ‘A lover of honour finds in himself no need to be known [for his merits]’ (LE, 27:665.6–7). Hence the two basic kinds of love of honour differ from each other with respect to their ground. Even though the physical love of honour can be unselfish, it must be classified under self-love in the broadest sense because it is based on inclination (and con-

³³ The analyses of this topic offered by Smit and Timmons (2015) and Denis (2014) are extremely helpful. However, neither Smit and Timmons nor Denis make an attempt to connect love of honour with love’s general framework – i.e. they do not problematise the status of love of honour as love. As with the animal impulses of self-love, Kant never discusses love of honour in terms of the general division of love. Given my previous analysis of love and desire, we can see why Kant might call the physical love of honour ‘love’: as an inclination, it is connected with the (lower) faculty of desire. There is also a long-running linguistic convention involved here: even Aristotle referred to desire for honour as ‘love of honour’ (1984; Nicomachean Ethics, Book IV, 1125b1–25; see Denis 2014, p. 203). On the other hand, as I will show below, moral love of honour comes close to respect.

³⁴ In particular, I’m speaking about the Collins notes on ethics.

³⁵ ‘Ein Ehrliebender findet in sich kein Bedürfnis, bekannt zu sein’.
nected with emulation). Moral love of honour, or ‘true love of honour’ [wahre Ehrliebe] (see LE, 27:695.26) would seem to be a kind of self-respect.

For the mature Kant, of course, animality cannot serve as the basis of morality, which must be grounded in pure practical reason. In the Vigilantius notes on ethics, however, true love of honour can come about when animal nature is educated by cultivating the natural, inborn tendency to emulation, such that one learns gradually to test one’s worth not against others but against the moral law (LE, 27:695.20–26). Kant does not elaborate on how one learns this, but it would probably be through moral education (see 9:446–449). My conceptual point is that by looking at love of honour, it is possible to detect a predisposition to the good in the animal self-love of sociality. Through the unsociably sociable self-love of humanity, animal sociality connects to love of honour and may thus be a contributing factor in the acknowledgment of the fundamental equality of all human beings as moral beings. But as Denis (2014, p. 205) warns us, maybe the continuity in Kant’s accounts of love of honour shouldn’t be exaggerated, and my basic claim in this section does not hinge on this tentative suggestion concerning the role of love of honour in mediating animality and morality.

The bulk of my discussion in this section has been motivated by a desire to understand. I have merely argued that Kant views the strongest impulses of human nature in terms of love, and that these impulses can be classified under animal mechanical self-love. Moreover, I am confident that animal mechanical self-love is somehow a predisposition to the good. The social impulse of animal self-love grounds the unsociably sociable self-love of humanity, which belongs to the history and teleology of moral progress. Love of honour is an interesting notion along this trajectory. At the very least, animal mechanical self-love is, for Kant, a necessary condition of the continued survival of our species.

1.2 The Middle Level of Self-Love

I shall now move on to what I call the broad middle level of self-love. The first section already identified this level from the perspective of the philosophy of history. Here, my focus will be moral philosophy. The moral perspective marks the self-love normally referred to in discussions of self-love in Kant, and it is clearly

36 The next section will show, with reference to the second Critique, that all inclination-based maxims fall under self-love, even if they are not ‘selfish’ in the sense of seeking one’s own advantage.
this kind of self-love that Kant had the most to say about. In order to make sense of this notion, I will consider the operation of self-love in the mature period, focusing on the *Groundwork*, the second *Critique*, and the *Religion*, while also taking into account the Vigilantius lectures on ethics and *The Metaphysics of Morals*.

I will approach this level of self-love with an interpretative key provided in the *Religion*, according to which love in general, and self-love in particular, is divided into two distinct forms or types of love, namely love of benevolence and love of delight. For this reason, I will begin by grounding my interpretation in the *Religion* and then work my way back chronologically through the *Groundwork* and the second *Critique*. We might call this the general division approach to the middle level of self-love.

Considering the overall aims of the present chapter and this study as a whole, there are three main claims that I shall argue for in this section: 1) Kant’s moral philosophy contains a notion of self-love that requires reason (and that is hence distinct from and cognitively ‘higher’ than animal mechanical self-love); 2) this self-love may be plausibly interpreted in terms of the general division of love, such that self-love is mainly benevolence toward oneself but also delight in oneself; and 3) while the latter especially brings with it the danger of self-conceit [*Eigendünkel*], which is evil, both forms of love are acceptable under moral conditions.

I take it that none of these claims are trivial, and they may be even more prone to objections than my assertions in the previous section. It would be possible to argue on the basis of certain passages in the *Groundwork* that animal mechanical self-love and ‘higher’ self-love are in fact not distinguishable from each other, and that the general division of love is simply a peculiarity of the *Religion*, and not supported by Kant’s other major works. However, I hope to show that these lines of argumentation will not work. The textual evidence suggests that Kant developed his account of self-love substantially after the *Groundwork*, and since the later works are fairly consistent in outline among themselves, their picture of self-love is preferable. This is not to say that Kant essentially changed his mind on self-love after 1785 or that the period following the *Groundwork* was marked by complete uniformity. I wish only to suggest that certain distinctions that we find in later works were not yet clearly in place in 1785. A further possible objection against my claims is that, even if it were granted that part of self-love involves reason and that the general division of love has something to do with it, this self-love still contains elements, or some meaning, that cannot be grasped in terms of the general division. At face value, this objection seems to me to be the most promising, and its success hinges on whether or not the general division is robust enough to capture Kant’s most all-encompassing formula-
tions of self-love, where self-love informs all instrumentally rational non-moral activity. I believe the general division is indeed robust at the middle level and that the general division of love, together with the multilevel account of self-love, makes it possible to harmonise the self-love discussed in the *Groundwork* with the later works while continuing to appreciate their apparent differences.

As mentioned above, in the fourth section of the first part of the *Religion* (‘Concerning the Origin of Evil in Human Nature’) Kant introduces the idea of a general division of love, which is applicable to self-love in particular: ‘Like love in general, *self-love* too can be divided into love of benevolence and love of delight (*BENEVOLENTIAE ET COMPLACENTIAE*), and both (as is self-evident) must be rational.’ (R, 6:45.22–25) The first amounts to wanting [*wollen*] things to go well [*wohl*] for oneself, and it is rational insofar as the instrumental reasoning related to the pursuit of happiness [*Glückseligkeit*] or well-being [*Wohlgefallen*] is apt. The rational maxims of love of benevolence for oneself are non-moral; they subordinate reason to natural inclination. (R, 6:45.25–32) Rational love of delight, on the other hand, has two senses, the first of which coincides with love of benevolence: ‘we take delight in those maxims, already mentioned, which have for the end the satisfaction of natural inclination [...] and then it is one and the same with love of benevolence toward oneself’ (R, 6:45.35–38). The second sense of love of delight for oneself is distinctive and seems to imply morality: ‘Only the maxim of self-love, of *unconditional delight* in oneself (independent of gain or loss resulting from action), is however the inner principle of contentment only possible for us on condition that our maxims are subordinated to the moral law.’ (R, 6:45.40–46.18)

This passage supports point 1) above: there is *some sort of* rational self-love in Kant that is distinct from the way animal mechanical self-love is described (this point is corroborated by the existence of the ‘self-love of humanity’ discussed in the latter half of the previous section; see R, 6:27.4–12). Hence, it may be proposed that there are at least two ‘levels’ of self-love in Kant: a non-

---

37 ‘Vom Ursprunge des Bösen in der menschlichen Natur.’
38 ‘Wie Liebe überhaupt, so kann auch Selbstliebe in die des Wohlwollens und des Wohlgefallens (*BENEVOLENTIAE ET COMPLACENTIAE*) eingetheilt werden, und beide müssen (wie sich von selbst versteht) vernünftig sein.’
40 ‘Allein die Maxime der Selbstliebe des unbedingten (nicht von Gewinn oder Verlust als den Folgen der Handlung abhängenden) Wohlgefallens an sich selbst würde das innere Prinzip einer allein unter der Bedingung der Unterordnung unserer Maximen unter das moralische Gesetz uns möglichen Zufriedenheit sein.’
rational level and a rational level. What I wish to focus on next is claim 2), which concerns the viability of the general division of love in this context. By showing that the general division of love is indeed viable, I hope to provide the necessary conceptual tools for establishing and clarifying the meaning of 3) – the claim that self-love is acceptable under moral conditions while self-conceit is evil.

In the *Groundwork*, self-love is explicitly contrasted with respect for the moral law. The latter represents a worth that ‘infringes’ [*Abbruch thut*] on self-love (GW, 4:401.28–29fn.) in the sense that we are subject to the moral law ‘without consulting self-love’ (GW, 4:401.33fn.)41. When explaining the proper motivational ground of moral action, Kant uses *Selbstliebe* as a general term to capture those motivational grounds that are distinct from morally adequate motivation. Pure morality abstracts from the presentation of ends, and by implication self-love can be seen to refer generally to cases where an expected ‘effect’ [*Wirkung*] of conduct is made the motivational ground of action. It seems to be self-love that informs all our fears and inclinations. (GW, 4:401) In this sense, in the *Groundwork* picture, self-love simply denotes all of our non-moral interests (see also GW, 4:406) – a picture that is consistent with the second *Critique*. Throughout his writings, Kant assumes that what human beings (as imperfect rational creatures) naturally desire is happiness, and hence self-love may be identified as the principle of subjective non-moral happiness. In fact, it is the name given to the subjective interest in happiness.

But the *Groundwork* does not clearly distinguish between rational and non-rational aspects of self-love. In Section I, ‘preservation’ [*Erhaltung*] and ‘prosperity’ [*Wohlergehen*] are identified as the constituents of happiness, which according to Kant is better pursued with instinct rather than reason. Here, self-preservation (which Kant later calls ‘love of life’) is not distinguished from prosperity in terms of reason, as is the case in the *Religion* (GW, 4:395.4–12; R, 6:26.12–18; cf. 6:45fn.). However, happiness as the satisfaction of the sum of all one’s inclinations (GW, 4:394.17–18; see Timmermann 2007, p. 20) in the *Groundwork* is a purpose that is assumed to be naturally necessary in all imperfect rational creatures (GW, 4:415.28–33). In other words, even though the distinction between animal mechanical self-love and rationally comparative self-love is not clear in the *Groundwork*, Kant says neither that self-love is non-rational or irrational nor that it is impermissible as such. We must interpret his idea such that, in the light of his basic division between nature and freedom/morality, self-love concerns our natural inclinations (it flows from the side of nature, as it were) but is entangled in our rational capacities.

41 ‘ohne die Selbstliebe zu befragen’.
In the *Groundwork*, the pursuit of happiness comes in the guise of ‘hypothetical assertoric imperatives’[^42], which means that the pursuit involves principles or quasi-commands (see GW, 4:414–415; cf. 4:418.28–32) stemming from means-end reasoning concerning the purpose of ‘one’s own greatest well-being’ (GW, 4:416.2)[^43]. In this limited and qualified sense, Herbert Paton (1947, p. 91) is correct: self-love in the *Groundwork* is ‘an objective principle of practical reason’, for it is assumed as actual in all finite rational beings. But the ‘objective’ here should not be confused with ‘objectively necessary’ in the moral sense. Self-love is not a duty, and as such it is not in the service of morality. Indeed, what is at issue in Kant’s examples of the specific duties not to commit suicide and not to make false promises is ‘the principle of self-love’ [*Princip der Selbstliebe*] (GW, 4:422.7; 4:422.24), but the key point is that the self-love in question may not be universalised (see Timmermann 2007, p. 81). In Kant’s examples of how the principle of self-love functions in moral life, what we encounter are immoral instances of self-love. Suppose, for instance, that I borrow money and falsely promise to repay it. This may well promote my ‘own benefit’ [*eigenen Zuträglichkeit*] and ‘future well-being’ [*künftigen Wohlbefinden*], but if universalised the promise would contradict itself, for then no one would believe others’ promises (GW, 4:422.15–36).[^44] Self-love may be objectively actual (its principle is active in all humans), but it is by no means objectively good (good ‘as such’ or ‘in itself’).

Can the natural but rational and principled self-love of the *Groundwork* be interpreted in terms of the general division of love found in the *Religion*? Instances of self-love in the *Groundwork* seem to refer to one’s own happiness or well-being, and in this sense self-love in the *Groundwork* may well be viewed along the lines of love of benevolence in the *Religion*. In short, it consists in wishing or willing that things will go well for oneself. There is no talk in the *Groundwork*

[^42]: According to Kant, a rational being is endowed with a will – that is, the capacity of acting from principles, which are representations of laws. Objective principles are rational commands, the formulae of which are imperative. Imperatives are further divided into hypothetical and categorical; hypothetical imperatives concern means-end reasoning, whereas only categorical imperatives are moral and do not refer to other ends. Hypothetical imperatives are further divided into problematic and assertoric practical principles, where the former concern possible purposes and the latter actual purposes. Moral imperatives are apodictically practical principles. (GW, 4:412–415) However, hypothetical imperatives are not commands strictly speaking but rather ‘counsels’ [*Anrathungen*] (GW, 4:418.31). In the second *Critique*, the hypothetical imperatives of self-love are, properly speaking, theoretical; they merely point out empirical causal connections and do not necessitate universally in determining the will. (C2, 5:25–26)

[^43]: ‘seinem eigenen größten Wohleyn’.

[^44]: I do not aim here to determine whether this argument is ultimately successful.
of any specific pleasure or delight taken in the maxims of self-love, which would seem to exclude the possibility of interpreting the self-love of the *Groundwork* directly in terms of love of delight. In the *Groundwork*, Kant writes that ‘skill in the choice of the means to one’s own greatest well-being can be called *prudence*’ (GW, 4:416.1–3)\(^{45}\), even though the subjective concept of happiness cannot be determined with complete certainty (GW, 4:418). In the *Religion*, love of benevolence is rational if it is ‘consistent with’ [zusammen bestehen] and ‘apt’ [tauglich] with regards to happiness at the level of one’s choices (R, 6:45.27–30). These rationality criteria seem to match the *Groundwork* definition of prudence quite closely. This supports the idea that while the non-rationality/rationality divide in the *Groundwork* is occasionally fuzzy when it comes to self-love, the principle of self-love discussed in the *Groundwork* is indeed love of benevolence. Hence, while the general division of self-love is not explicitly operative in the *Groundwork*, the principle of self-love may be plausibly interpreted in terms of it, which lends support to claim 2) above.\(^{46}\)

I now move on to the second *Critique*. For explanatory purposes, we can roughly divide Kant’s treatment of self-love in this work into two ‘phases’ according to the *function* of self-love in the discussion. It seems to me that the function of self-love in the first chapter of the ‘Analytic’ is more or less equivalent to how self-love operates in the *Groundwork*: according to Kant’s dualistic conception of possible determining grounds of the will, self-love functions as a general term for the determining ground of a will not determined by the moral law (see fig. 1). In the third chapter, on the other hand, Kant discusses how the moral law ought to influence the will *as regards self-love* and the moral danger that ensues if this

---

\(^{45}\) ‘kann man die Geschicklichkeit in der Wahl der Mittel zu seinem eigenen größten Wohlsein *Klugheit* [...] nennen.’ Note that in this passage Kant speaks of prudence ‘in the narrowest sense’ [im engsten Verstande]. It is unclear what the broader sense would be, but Kant explains in a footnote that prudence divides into ‘worldly prudence’ [Weltklugheit] and ‘private prudence’ [Privatklugheit]. The former is the skill of using others for one’s purposes; the latter, the unification of one’s purposes to one’s ‘enduring advantage’ [daurenden Vortheil] (GW, 4:416.30–33). I take it that private prudence is the broader of the two, but this is in tension with the equally plausible notion that skill or insight [Einsicht] to unite one’s purposes belongs to skill in choosing the means to one’s ‘own greatest well-being’ (which is prudence in the narrowest sense). (Cf. Kain 2003, pp. 247, 263fn.60)

\(^{46}\) Note that even if my interpretation of self-love in the *Groundwork* is rejected, we might still need to accept the general division interpretation of the middle level of self-love on the basis of the relative harmony between the second *Critique* and the *Religion*. In this case, we would need a weighty argument as to why the doctrine of self-love in the *Groundwork* overrides the combination of the second *Critique* and the *Religion*. However, it is difficult for me to see how this could be charitable, and I don’t have such an argument in sight.
does not take place properly (the problem of self-conceit). I will now consider both of these ‘phases’ in the light of the general division.

The lion’s share of the first chapter of the second Critique is dedicated to establishing that the practical principles of self-love (which Kant calls ‘material’) are not the principles generated by the moral law (which is ‘formal’). The basic idea is summarised in Kant’s famous statement: ‘All material practical principles as such are, without exception, of one and the same kind and come under the general principle of self-love or one’s own happiness.’ (C2, 5:22.6–8) Equally telling is the following: ‘The direct opposite of the principle of morality is the principle of one’s own happiness made the determining ground of the will’ (C2, 5:35.7–8). We learn that in the case of self-love, the determining ground of the will always rests on expectations of pleasure or displeasure, ‘agreeableness’ [Annehmlichkeit] or ‘disagreeableness’ [Unannehmlichkeit]. The fact that pleasure is involved makes it plausible to suppose that the self-love at issue might be love of delight. This conjecture is supported by the fact that

---

47 As Jens Timmermann has pointed out to me, the discussion is also directed against rival ethical schools, most notably the hedonism of Epicurus (see C2, 5:24.15–20).

48 ‘Alle materiale praktische Principien sind, als solche, insgesammt von einer und derselben Art und gehören unter das allgemeine Princip der Selbstliebe oder eigenen Glückseligkeit.’ It is important to note that the broad formulation of self-love implies that even beneficent love of neighbour, if based on inclination, is technically subsumed under self-love. However, this does not mean that loving one’s neighbour out of inclination is selfish; it merely means that in the choice of the objects of beneficent love from inclination we are prone to a ‘second order self-partiality’ (see Wood 1999, p. 271). Note also that Kant occasionally distinguishes between sympathy [Sympathie] and self-love [Philautie] (C2, 5:85.12).

49 ‘Das gerade Widerspiel des Princips der Sittlichkeit ist: wenn das der eigenen Glückseligkeit zum Bestimmungsgrunde des Willens gemacht wird’.
there are no instances of ‘benevolence’ [Wohlwollen] in the first chapter of the ‘Analytic’. With this said, is ‘delight’ [Wohlgefallen] mentioned in this context? As Gregor notes, ‘Kant [...] uses a variety of words for “pleasure”’ (C2, p. 157fn.a), and words such as Lust, Vergnügen, Zufriedenheit and Wohlgefallen are all quite close to each other. As determining grounds of the will, they belong to the ‘lower faculty of desire’ [unteres Begehungsvermögen], as opposed to the ‘higher faculty’ [oberes Begehungsvermögen], which alone is moral and furnished by pure practical reason irrespective of sensibility. And indeed, under Remark I of the basic theorem of self-love (Theorem II), we find a chain of signifiers that goes from ‘satisfaction’ [Vergnügen], ‘joys’ [Freuden], ‘amusements’ [Ergötzungen] and ‘enjoyment’ [Genuß] back to ‘satisfaction’ [Vergnügen], and finally the word Wohlgefallen appears (C2, 5:24.11–12). The point is that even though the pleasures we take in ‘overcoming obstacles opposed to our plans’ (C2, 5:24.4–5) are ‘more refined because they are more in our control than others’ (C2, 5:24.6–7), the determining ground of this delight is nevertheless our own happiness, or the lower faculty of desire – i.e. self-love.

It thus seems plausible to identify the self-love at issue in Chapter I of the ‘Analytic’ as love of delight. But does this create an inconsistency with the Groundwork, where I identified self-love as love of benevolence? In both cases, the basic point is clearly that self-love and the moral law constitute two opposing grounds of action or determination of the will. How is this apparent similarity regarding the function of self-love to be explained if we are dealing with two different dimensions of the general division of love? Quite easily, I think, if we recall that Kant explains in the Religion that the first sense of love of delight is in fact identical to love of benevolence (R, 6:45fn., see above). In this case, love of delight simply denotes taking pleasure in the inclination-based maxims of love of benevolence (insofar as these maxims yield successful results) ‘like a mer-

---

50 I here follow Timmermann’s suggestion (private discussion) when it comes to translating Ergötzungen. Gregor uses ‘delights’, but this would be confusing as I render Wohlgefallen ‘delight’. In this context, Gregor has Wohlgefallen as ‘pleasure’.

51 ‘Überwindung der Hindernisse, die sich unserem Vorsatze entgegensetzen’.

52 ‘feinere [...] , weil sie mehr wie andere in unserer Gewalt sind.’

53 Cf. Christine Korsgaard (1997, p. 220), who argues, in a supposedly Kantian fashion, that the instrumental reason related to pursuing our aims shares ‘a common normative source’ with moral reason in the autonomy of the agent. Even though specifically moral aims also often require auxiliary instrumental reasoning, I think that Korsgaard’s account, at least from an exegetical perspective, does not sufficiently appreciate Kant’s basic distinction between self-love and the moral law as two opposing sources of motivation for action, which he so vehemently holds in both the Groundwork and the second Critique. Korsgaard’s views have been forcefully criticised by Camilla Kong (2012).
chant whose business speculations turn out well and who, because of the max-
ims he adopted therein, rejoices in his good insight’ (R, 6:45fn.)⁵⁴. The fact that
Kant identifies this taking of pleasure with ‘wanting things to go well for oneself’
implies that the notion of pleasure is implicitly included in the love of self-be-
nevolence as such. The general division therefore helps to make clear that, at
a general level, the self-love of the first chapter of the second Critique is not
only consistent with the self-love of the Groundwork but can also be viewed as
an elaboration of it.⁵⁵

However, in the ‘second phase’ of Kant’s discussion of self-love in the sec-
ond Critique, things become exegetically more complicated. In Chapter III, ‘On
the Incentives of Pure Practical Reason’ (C2, 5:72.27)⁵⁶, Kant discusses the effects
of the moral law on the mind as an incentive or ‘spring’ [Triebfeder] (see Eng-
strom 2010, pp. 91–93). Because Kant’s basic system of the mental structure of
the agent relies on the dualism of the moral law and sensibility, the effects of
the moral law must be felt as effects on sensibility. We know that allowing the
sensuous side of our nature to determine the will is equated with self-love in
a broad sense, and thus the effects of the moral law are necessarily effects on
self-love (inclinations, feeling). A will determined by the moral law must reject
all such self-love that is opposed to the law. In this context, Kant refines the con-
ceptual framework of self-love, introducing new systematic distinctions:

All the inclinations together (which can be brought into a tolerable system and the satisfac-
tion of which is then called one’s own happiness) constitute regard for oneself (SOLIPS-
MUS). This is either the self-regard of love for oneself, a predominant benevolence toward
oneself (PHILAUTIA), or that of delight in oneself (ARROGANTIA). The former is called,
in particular, self-love, the latter, self-conceit. Pure practical reason merely infringes upon
self-love, inasmuch as it only restricts it, as natural and active in us even prior to the
moral law, to the condition of agreement with this law, and then it is called rational self-
love. But it strikes down self-conceit altogether [...]. (ibid., 5:73.9–18)⁵⁷

⁵⁴ ‘wie ein Kaufmann, dem seine Handlungsspeculationen gut einschlagen, und der sich wegen
der dabei genommenen Maximen seiner guten Einsicht erfreut.’ Pluhar’s translation.
⁵⁵ I am not claiming that the two accounts are technically identical. For instance, there are dif-
fferences in the way Kant construes his treatment of hypothetical imperatives, but I cannot dis-
cuss these differences in detail here.
⁵⁶ ‘Von den Triebfedern der reinen praktischen Vernunft’.
⁵⁷ ‘Alle Neigungen zusammen (die auch wohl in ein erträgliches System gebracht werden kön-
nen, und deren Befriedigung alsdann eigene Glückseligkeit heißt) machen die Selbstsucht (SOL-
IPSISMUS) aus. Diese ist entweder die der Selbstliebe, eines über alles gehenden Wohlwollens
gegen sich selbst (PHILAUTIA), oder die des Wohlgefallens an sich selbst (ARROGANTIA). Jene
heißt besonders Eigenliebe, diese Eigendünkel. Die reine praktische Vernunft thut der Eigenliebe
blos Abbruch, indem sie solche, als natürlich und noch vor dem moralischen Gesetze in uns
Is the general division of love into love of benevolence and love of delight operative in this passage? Yes and no. What Kant generally calls self-love in other parts of the *Groundwork* and the second *Critique* is now called self-regard [*Selbstsucht*]. Here, Kant seems to use the term ‘self-love’ [*Selbstliebe, Eigenliebe*] in a more restrictive sense: self-love ‘in particular’ [*besonders*] is benevolence towards oneself, and it is distinguished from delight in oneself, which is ‘self-conceit’ [*Eigendünkel*]⁵⁸. The basic notions of the general division of love, benevolence and delight, are explicitly operative, but ‘self-love’ is here connected only to benevolence. This suggests that love of benevolence is a more paradigmatic form of self-love. But what should we make of the fact that delight is discussed here not in terms of love but in terms of self-conceit, which, as implied by the context, is obviously reprehensible? Does this cast doubt on the ‘acceptability-thesis’ of claim 3 above, according to which both benevolence for oneself and delight in oneself are somehow acceptable? Or does it mean that the delight in oneself that is under discussion here is not part of self-love? And if delight as self-conceit does not belong to self-love, what are the implications for the success of the general division approach in the particular context of self-love?

I will address from three different directions the problem of how self-love, delight in oneself, and self-conceit interrelate, which will clarify the status of the general division of love in this context and bring to a close my discussion of the middle level of self-love. These three directions are 1) the generality of self-love, 2) the problem of rationality in self-love, and 3) the Vigilantius notes.

1) Firstly, there is a sense internal to the second *Critique* in which both benevolence and delight in Chapter III can be, or even must be, interpreted as belonging to self-love in the broadest middle-level sense. Kant’s description of self-regard as the sum of inclinations (as a tolerable system) necessarily falls under the basic description of self-love in ‘Theorem II’, where self-love includes ‘all material practical principles’ (C2, 5:22.6)⁵⁹. We can understand the relation between self-regard and self-love (in the broad sense of the second *Critique*) either such that they are equivalent, insofar as the system of self-love is ‘tolerable’ [*erträglich*], or such that self-regard is a name used for self-love when everything that rege, nur auf die Bedingung der Einstimmung mit diesem Gesetze einschränkt; da sie alsdann *vernünftige Selbstliebe* genannt wird. Aber den Eigendünkel schlägt sie gar nieder*.  
⁵⁸ For an earlier, pre-autonomy treatment of self-love and self-conceit, see the Collins notes on ethics, where self-love or *philautia* is described as love of delight and contrasted with self-conceit (LE, 27:357).  
⁵⁹ ‘Alle materiale praktische Principien’.
falls under self-love is *taken together* (see above, 5:73.8–19). In both cases, any instance of benevolence for oneself or delight in oneself is subsumed under self-love (irrespective of whether the system is ‘tolerable’). Further along in Chapter III Kant explains that self-love is the propensity to turn ‘subjective determining grounds of choice into the objective determining ground of the will in general’, whereas ‘if self-love makes itself lawgiving, it can be called *self-conceit.*’ (C2, 5:74.15–19)60 The latter passage confirms that self-conceit belongs to self-love: it is self-love that arrogantly takes the place of the moral law.61 Thus the fact that Kant distinguishes between self-love and self-conceit cannot as such be taken to warrant the interpretation that the latter does not belong to the former. The viability of the general division approach is not threatened on the mere basis of the existence of such a division.

2) In the light of the second Critique and the Religion, we now have at least three different senses of ‘rational self-love’, which may or may not turn out to be equivalent. In the Religion, love of benevolence as willing one’s own happiness is ‘rational to the extent that with respect to the end only what is consistent with the greatest and most abiding well-being is chosen, and that also the most apt means for each of these components of happiness are chosen.’ (R, 6:45.27–30)62 In the second Critique, love of benevolence is rational when it is ‘infringed’ by the moral law ‘to the condition of agreement with this law’ (C2, 5:73.15–17).63 If we interpret consistency ‘with the greatest and most abiding well-being’ as implying that one does not act immorally, then the two senses of rational love of benevolence in the Religion and the second Critique can indeed be viewed as equivalent. This would imply that they cohere with the prudence discussed in the *Groundwork* and that Kantian prudence as such implies

---

60 ‘den subjectiven Bestimmungsgründen seiner Willkür zum objectiven Bestimmungsgrunde des Willens überhaupt’ / ‘welche [die Selbstliebe], wenn sie sich gesetzgebend und zum unbedingten praktischen Princip macht, *Eigendünkel* heißen kann.’

61 The self-love that serves as the ground of self-conceit can now be viewed as either love of benevolence toward oneself (wanting things to go well for oneself) or love of delight as taking pleasure in the maxims of the former (which would amount to the same thing as the former on the basis of the Religion). For detailed discussions of self-conceit, see Reath (2006, pp. 14–17, 23–25), Engstrom (2010), and Moran (2014).

62 ‘vernünftig, als theils in Ansehung des Zwecks nur dasjenige, was mit dem größten und dauerhaftesten Wohlergehen zusammen bestehen kann, theils zu jedem dieser Bestandstücke der Glückseligkeit die tauglichsten Mittel gewählt werden.’

63 ‘Abbruch thut’ / ‘auf die Bedingung der Einstimmung mit diesem Gesetze’.
a moral side constraint: the prudent agent does not act immorally.\textsuperscript{64} This interpretation is charitable, for it renders the mature period consistent in this respect. But the interpretation is not necessary, and the text remains ambiguous. It is equally possible that there are two different rationality criteria for love of benevolence, such that rationality implies mere instrumentality irrespective of morality in the case of the \textit{Religion} (and, supposedly, the \textit{Groundwork}) and the ‘stronger’, morally constrained instrumental rationality appears only in the second \textit{Critique}. Fortunately, my argument does not hinge on this point. The third type of rational self-love concerns love of delight (as distinct from love of benevolence). In the \textit{Religion}, the condition of rationality for this self-love is quite simply morality in the sense in which morality was first opposed to self-love in the \textit{Groundwork} and the second \textit{Critique}. For this reason, Kant has doubts about the appropriateness of the term ‘rational self-love’ in this case:

We could call this love a \textit{rational love} of oneself that prevents any adulteration of the incentives of the power of choice by other causes of contentment consequent upon one’s actions (under the name of happiness to be procured through them). But, since this denotes unconditional respect for the law, why needlessly render more difficult the clear understanding of the principle with the expression \textit{rational self-love}, when this self-love is however \textit{moral} only under the latter condition, and we thus go around in a circle (for we can love ourselves morally only to the extent that we are conscious of our maxim to make respect for the law the highest incentive of our power of choice)? (R, 6:46.21–30)\textsuperscript{65}

Is this the same sense in which rational love of benevolence is described as ‘rational’ in the second \textit{Critique}? Clearly not, since the rationality criterion for rational love of benevolence in the second \textit{Critique} is morality merely in the negative sense: reason may still be used in the service of inclination, and it suffices that the prudential maxim is not contra-moral. Yet the rational moral love of delight in the \textit{Religion} implies a contentment [\textit{Zufriedenheit}] that is only possible through ‘unconditional respect for the law’. It is also worth noting that in the second \textit{Critique} passage quoted above (C2, 5:73.9–18), ‘self-love’ is not mentioned in

\textsuperscript{64} Note, however, that even in this case there is nothing moral about the self-love as such. It is just negatively constrained by morality.

\textsuperscript{65} ‘Man könnte diese die \textit{Vernunftliebe} seiner selbst nennen, welche alle Vermischung anderer Ursachen der Zufriedenheit aus den Folgen seiner Handlungen (unter dem Namen einer dadurch sich zu verschaffenden Glückseligkeit) mit den Triebfedern der Willkür verhindert. Da nun das letztere die unbedingte Achtung fürs Gesetz bezeichnet, warum will man durch den Ausdruck einer \textit{vernünftigen}, aber nur unter der letzteren Bedingung \textit{moralischen Selbstliebe} sich das deutliche Verstehen des Princips unnöthigerweise erschweren, indem man sich im Zirkel herumdreht (denn man kann sich nur auf moralische Art selbst lieben, sofern man sich seiner Maxime bewußt ist, die Achtung fürs Gesetz zur höchsten Triebfeder seiner Willkür zu machen)?’
the context of delight in oneself, which is defined as self-conceit. Does this mean that the doctrines of self-love in the second *Critique* and the *Religion* are incompatible, or at least that they cannot be plausibly harmonised in terms of the general division of love? No, for in the second *Critique* we do find a notion of acceptable delight in oneself that is not self-conceit and that appears to coincide with the description of moral rational love of delight in the *Religion*. In the section on the ‘Critical Resolution of the Antinomy of Practical Reason’ (C2, 5:114)\(^\text{66}\), Kant introduces the notion of a satisfaction or delight [*Wohlgefallen*] in oneself that is not based on material determining grounds and instead rests solely on the direct determination of the will by the moral law. (C2, 5:116.25–33) This is ‘contentment with oneself’ [*Selbstzufriedenheit*], which is ‘a negative delight in one’s existence, in which one is conscious of needing nothing.’ (C2, 5:117.28–31)\(^\text{67}\) It denotes ‘mastery over one’s inclinations’ [*der Obermacht über seine Neigungen*] rather than ‘complete independence’ [*gänzliche Unabhängigkeit*] from them. (C2, 5:118.26–33) I interpret its negativity to mean that it is a delight that is not based on the positive satisfaction of any inclination (unlike the first sense of love of delight in the *Religion*) and merely follows from consciousness of being able to ‘negate’ inclinations through respect for the moral law. There are therefore two kinds of delight in oneself in the second *Critique*: self-conceit [*Eigendünkel*] and contentment with oneself [*Selbstzufriedenheit*]. The former (self-love as the unconditional practical principle in C2, 5:74.17–19) is practically identical with the *Religion*’s idea of self-love as the spring of evil: ‘self-love [...], when adopted as the principle of all our maxims, is precisely the source of all evil.’ (R, 6:45.14–15)\(^\text{68}\) Self-contentment in the second *Critique* comes close to the second sense of love of delight in the *Religion*, which is also contentment [*Zufriedenheit*] with oneself through respect for the moral law. Hence, while the conceptual systems of self-love in the second *Critique* and the *Religion* seem different, these differences can be considered superficial. The respective notions of self-love can be harmonised, which provides support for claims 2) and 3) above concerning the

---

\(^\text{66}\) ‘Kritische Aufhebung der Antinomie der praktischen Vernunft’.

\(^\text{67}\) ‘ein negatives Wohlgefallen an seiner Existenz andeutet, in welchem man nichts zu bedürfen sich bewußt ist.’

\(^\text{68}\) ‘Selbstliebe [...], als Princip aller unserer Maximen angenommen, gerade die Quelle alles Bösen ist.’ It now seems to me that the reason self-conceit is not discussed in terms of rational self-love either in the second *Critique* or in the *Religion* is that while the notion of self-conceit implies the existence of the rational capacity (its possibility belongs to the predisposition to humanity), it involves a morally improper use of reason, and in this sense it is *irrational*. Recall that the ground of evil cannot be located in a natural impulse and that its rational origin remains inscrutable (R, 6:32–43).
middle level of self-love: the general division of love is in operation, and both
love of benevolence for oneself and love of delight in oneself are acceptable
under moral conditions.\textsuperscript{69}

3) The Vigilantius notes on ethics from 1793–94 corroborate the picture of the
middle level of self-love I have drawn thus far. They explicitly confirm that
Kant’s use of the general division is not simply an anomaly found only in the
\textit{Religion} but is rather foundational to how Kant viewed self-love in the mature
period. They also explain the relation between the conceptual systems of the sec-
ond \textit{Critique} and the \textit{Religion}. In the Vigilantius notes, self-love is discussed as
\textit{philautia}. \textit{Philautia} is divided into love of benevolence and love of delight. It
is implied that love of benevolence means the inclination or the will to promote
one’s own ends, and there is no reason to view the account as differing signifi-
cantly from Kant’s previous discussions.\textsuperscript{70} Love of benevolence is assumed to be
active in all human beings, and it is acceptable. If the duty of love towards oth-
ers\textsuperscript{71} is excluded from consideration, love of benevolence is ‘solipsism’ [\textit{Solipsismus}]\textsuperscript{72}
or ‘egotism’ [\textit{Eigennutz}], in which case it becomes a moral flaw. (LE,
27:620) Love of delight, on the other hand, is more complex, and as we would
assume on the basis of the combination of the second \textit{Critique} and the \textit{Religion},

\textsuperscript{69} In general, the \textit{Religion} asserts that both love of benevolence and love of delight ‘must be
rational’. This should now be understood in both a descriptive and a normative sense: these
forms of self-love are rational on their own as a matter of definitional necessity (they imply
the level of reason), and they appear under a normative requirement to be harnessed by reason
or \textit{made rational}. Even if love of benevolence is construed as being completely non-moral, the
demand of aptness in instrumental reasoning imposes quasi-normative constraints on the oper-
ation of self-love. Love of delight may be moral as a consequence of the operation of pure prac-
tical reason (see also Section III of my Introduction).

\textsuperscript{70} In comparison with the second \textit{Critique} and the \textit{Religion}, the main shift of emphasis in Kant’s
discussion of self-love in the Vigilantius notes (and \textit{The Metaphysics of Morals}) is that whereas
self-love (as love of benevolence) was previously mainly about the maxims of actions and want-
ing things to go well for oneself, in the latter two sources \textit{philautia} is mainly about regarding
oneself as worthy of being loved by others [\textit{Liebenswürdig}]. Arrogance, on the other hand, is
an unwarranted claim to respect, and I believe at least part of the explanation for this shift
of emphasis has to do with the fact that the distinction between love and respect has such a
prominent place in the ‘Doctrine of Virtue’. Since self-love (as willing that things go well for
me or that others love me) brings about self-conceit if made an unconditional law, it seems
that somehow the will to be happy and loved by others brings about unjustified claims to respect
from others if universalised absolutely (see LE, 27:621).

\textsuperscript{71} See ch. 4.2.2.

\textsuperscript{72} Note that the notion of \textit{Solipsismus} seems more restrictive and negative in tone here than in
the second \textit{Critique} (see C2, 5:73).
three meanings of love of delight can be detected in Kant’s treatment. Even though love of delight forms the other half of the division of philautia in these notes, it seems that Kant reserves the paradigmatic meaning of philautia for love of benevolence.  

The first meaning, in which love of delight may be considered equivalent to love of benevolence, is skinned over in passing in the very first sentence, the latter part of which already implies the danger of irrational self-conceitedness: 1) ‘This, too, is philautia, if it is exclusively entertained towards oneself, but also becomes unreasonable’ (LE, 27:621.18 – 19). We then learn that in comparison with others, love of delight becomes ‘self-estimation of oneself’ [Selbstschätzung seiner selbst], which has two possible outcomes: 2) ‘arrogance’ [Arroganz] and 3) ‘true self-esteem’ [wahre Selbstschätzung]. (LE, 27:621.26 – 622.5) The first is a faulty, unjustified claim to respect from others, and given that in the second Critique Kant defined self-conceited delight precisely as arrogance, we may take it that this is the same delight that was at issue in Kant’s previous discussion of self-conceit. This arrogant delight is now discussed explicitly in terms of love of delight. True self-esteem, on the other hand, is based on ‘a close examination of oneself’ (LE, 27:621.29) and is only warranted on the condition of performing imperfect meritorious duties. As a morally conditioned and permissible delight in oneself, this corresponds roughly to the moral love of delight in the Religion (which implies unconditional respect for the moral law) and the negative moral delight of the second Critique (which implies moral freedom or mastery over inclinations). In Vigilantius, however, the emphasis shifts from negative to positive: ‘we thereby add a supplement to morality’ and ‘can thus acquire merit in relation to others’ (LE, 27:622.17 – 20). There is a positive moral delight in oneself in Vigilantius, which seems very close to the concept of ‘ethical reward’ in the actual Metaphysics of Morals, a reward which is similarly occasioned by the performance of wide duties and is ‘a moral pleasure that goes beyond mere contentment with oneself (which can

---

73 This was the case in the second Critique, where philautia was love of benevolence and distinct from self-conceited delight (C2, 5:73). This is also the case in The Metaphysics of Morals, where philautia is contrasted with self-conceit (MM, 6:462.5 – 10). The Vigilantius notes, however, allow for the construal of false claims of moral merit in Wohlgefährliche Selbstliebe as philautia (LE, 27:624 – 625). The context would seem to imply that this love of delight for oneself is in fact the self-conceit of the published works, but my argument does not hinge on this, and I do not wish to force the interpretation.

74 ‘Auch diese ist Philautie, wenn sie ausschließlich gegen sich selbst ausgeübt wird, wird aber auch vernunftwidrig.’

75 ‘vorgängiger genauere Prüfung seiner selbst’.

76 ‘wir geben dadurch der Moralität einen Zusatz’ / ‘Verdienste können wir indeß nur dadurch gegen andere enlangen’.
be merely negative)’ (MM, 6:391.12–13). Based on the combination of these passages, the interpretation must be that there are actually two possible forms of morally acceptable love of delight in oneself: negative self-contentment and positive moral pleasure from merit (both of which, however, involve the danger of self-conceit).

Let us take stock of what we’ve learned about the middle level of self-love thus far. I hope to have shown: 1) that there is a self-love that involves reason (one that is ‘higher’ than mere animal nature); 2) that even though the conceptual system of this level of self-love varies throughout the mature period, these variations can be made sense of by using the general division of love in a way that renders the outlines and main tenets of Kant’s discussion (relatively) consistent; and 3) that self-love is, paradigmatically, an acceptable love of benevolence for oneself common to all human beings, but it is also love of delight for oneself. I have also shown that love of delight may be 1) simply equivalent to love of benevolence, or 2) an acceptable moral self-contentment (negative) or self-esteem in a feeling of merit (positive), but it may also be 3) self-love turned into morally reprehensible self-conceit or arrogance. On this basis, the structural outline of the middle level of self-love may be presented in the form of a diagram (see fig. 2).

### 1.3 Self-Love and the Highest Good

We saw above that the only kind of self-love that Kant explicitly rejects is self-conceit or arrogance. Self-love that is made an unconditional law is ‘the source of all evil’, and the moral law ‘strikes down’ such self-conceit. But we have also seen that self-love is scarcely reducible to self-conceit, and that there are various

---

77 ‘einer moralischen Lust, die über die bloße Zufriedenheit mit sich selbst (die blos negativ sein kann) hinaus geht’.

78 In relation to others, the positive pleasure might actually be derived from a sympathetic reaction to the recipient of the moral action whose happiness is promoted by the virtuous act (see 6:391.16–21). I thank Jens Timmermann for making me aware of this (in private discussion). However, Timmermann’s line nevertheless leads to the question of why, then, Kant uses the term ‘moral pleasure’ [moralische Lust], if this pleasure is based on sensuous sympathy. One clue is that Kant is speaking of moral pleasure here in terms of ‘receptivity’ [Empfänglichkeit] (MM, 6:391.11). Moral pleasure might therefore be connected to Kant’s discussion of ‘moral feeling’ [das moralische Gefühl] in Section XII of the Introduction to the ‘Doctrine of Virtue’, where the natural predisposition of moral feeling is described as Empfänglichkeit to moral pleasure and displeasure (MM, 6:399.19; 6:400.18). Perhaps the moral pleasure here is not a sympathetic reaction but a positive dimension of ‘moral feeling’.
kinds of self-love that Kant views as necessary elements of our animal nature (animal mechanical self-love) and our humanity (love of benevolence). Arguably, there are also forms of self-love that Kant accepts on the basis that they are produced by morality (negative and positive moral love of delight). It must therefore be said that any strong interpretative claim made in direct opposition to self-love, such as Wood’s assertion that ‘love is not an attitude that clear-sighted and rational people could ever take toward themselves’ (Wood 1996, p. 157), is very difficult to sustain.\footnote{Towards the end of his article, Wood also suggests that ‘Kant holds that people of virtuous disposition will limit their claims on happiness to those of reasonable self-love.’ (Wood 1996, p. 157) However, Wood does not explain why reasonable self-love should not be considered an attitude of love towards oneself, and for this reason I admit to finding the main claim of his argument ambiguous.}

But isn’t there a demand in Kant to the effect that we should aim to mitigate or diminish our self-love? As a determining ground of the will, self-love is clearly opposed to the moral law, and it is equally obvious that we should aim to further and strengthen the position of the moral law as an incentive to our actions. When speaking of ‘love for the law’\footnote{I discuss love for the law in detail in ch. 3.1.} in the second Critique, Kant proposes that we ought to strive for a moral disposition such that we have no inclinations that are in tension with what morality commands. This would be a state of subjective moral perfection, where one’s attitude towards the moral law would be
not respect but actually love or liking [Zuneigung]. (C2, 5:83–84) Insofar as self-love is a general name for inclinations and is opposed to the moral law as a determining ground of the will, with love for the law self-love in this sense would cease to exist, and love for the law seems to be something that is called for. Similarly, the Religion asserts that what is required of us is a dispositional ‘revolution’ [Revolution] or a ‘change of heart’ [Änderung des Herzens], which consists in adopting the moral law as our supreme maxim and a gradual and laborious reworking of our conduct on this basis (R, 6:47–48).

But does this mean that we ought to get rid of self-love? Or in other words, is there or ought there to be an endpoint of human moral progress in which self-love is no longer a part of the constitution of our cognitive framework? I don’t think so. The second Critique discussion of love for the law makes very clear that because we are imperfect creatures, love for the law can never actually be achieved. We ‘can never be altogether free from desires and inclinations’, and love for the law remains a ‘constant though unattainable goal of [...] striving’ (C2, 5:84.4–15)⁸¹. It is of course the case that from a decidedly moral perspective Kant expresses as strong dislike of inclinations. In the Groundwork, inclinations ‘are so far from having an absolute worth [...] that to be entirely free from them must rather be the universal wish of every rational being’ (GW, 4:428.16–17)⁸². In the second Critique, inclinations ‘are always burdensome to a rational being’ who wishes ‘to be rid of them’ (C2, 5:118.7–9)⁸³. These statements are not surprising, since the basic business of pure practical reason is to disregard inclinations altogether when determining the will (see e.g. C2, 5:118.16–18). As already mentioned, however, even though inclinations can be disregarded in moral decision-making, Kant holds that ridding ourselves of inclinations entirely is not possible. When he considers the human being more broadly – as both a moral-rational and a natural-physical creature – his tone with regard to inclinations is different. In the Religion, he holds not only that we cannot rid ourselves of inclinations but also that we should not: ‘Considered in themselves natural inclinations are good, i.e. not reprehensible, and to extirpate them would not only be futile but harmful and blameworthy as well; we must rather only curb them’ (R, 6:58.1–4)⁸⁴. He clearly does not reject inclina-

⁸¹ ‘niemals von Begierden und Neigungen ganz frei sein’ / ‘beständigen, obgleich unerreichbaren Ziele [...] Bestrebung’.
⁸² ‘haben so wenig einen absoluten Werth [...] daß vielmehr, gänzlich davon frey zu seyn, der allgemeine Wunsch eines jeden vernünftigen Wesens seyn muß.’
⁸³ ‘sind sie einem vernünftigen Wesen jederzeit lästig’ / ‘ihrer entledigt zu sein’.
⁸⁴ ‘Natürliche Neigungen sind, an sich selbst betrachtet, gut, d.i. unverwerflich, und es ist nicht allein vergeblich, sondern es wäre auch schädlich und tadelhaft, sie auszotten zu wollen; man
tions or self-love as such. Kant understands full well that although we have a moral vocation, we are also natural creatures. Even though all inclinations lack moral worth and any inclination can be contra-moral at the token level, some inclinations, such as sexual inclination, can be viewed as predisposed to the good at least in the sense of being necessary for the continued survival of the species. Some inclinations can even be beautiful results of moral action (such as the benevolent inclination of love towards others that arises from being beneficent to others from duty) (C2, 5:82.18–20; MM, 6:402; see ch. 4 below). I therefore think that aiming for 'love for the law' does not constitute a demand not to have inclinations at all; rather, it involves strengthening the moral incentive within us so that inclinations have less power over us. This moral striving also involves the cultivation of at least some inclinations (such as the inclination of benevolence for others), to the end that our emotional dispositions harmonise with moral demands (see ch. 4.2.2.A). Striving for a morally better character therefore means making the moral law the supreme incentive in the sense that it is capable of overriding self-love.

There is thus no real demand in Kant’s moral philosophy to eradicate self-love from our lives. The most precise formulations of the thought that self-love should be limited or conditioned can be found in the second Critique and The Metaphysics of Morals. The way in which it is permissible to make self-love or love of benevolence lawlike for ourselves is by limiting the maxim of our own happiness with the happiness of others. We can call this the universalisability condition of the maxim of one’s own happiness. My own happiness ‘can become an objective practical law only if I include in it the happiness of others.’ (C2, 5:34.31–32) In this case, as we learn from the second Critique, self-love may acquire ‘objective validity’ [objective Gültigkeit], as the determining ground of the will is not self-love but the lawful form, through which self-love is limited and the notion of happiness is applied to all finite beings. In The Metaphysics of Morals, we find a mirror image of the same idea, now concerning ‘practical love’, or the grounding of the duty of active rational benevolence toward others [Liebe des

---

85 I thank Jens Timmermann for reminding me of this. Consider, for example, a father who (out of inclination-based benevolent love) rushes to buy a birthday present for his child and on his way to the toy store neglects to help someone who is drowning in a nearby pond.

86 This general view is inspired by Baron (1995).

87 ‘kann nur als dann ein objectives praktisches Gesetz werden, wenn ich anderer ihre in dieselbe mit einschließe.’
Wohlwollens]. Assuming that self-love is a universal aspect of humanity, it is now
the case that in order for there to be a duty to love others practically, one must
include oneself in the universalised maxim of love as an object of love:

I want everyone else to be benevolent toward me (BENEVOLENTIA); hence I ought also to
be benevolent toward everyone else. But since all others with the exception of myself would
not be all, so that the maxim would not have within it the universality of a law, which is still
necessary for imposing obligation, the law making benevolence a duty will include myself,
as an object of benevolence, in the command of practical reason. (MM, 6:451.4–10)

Kant explains that the above does not make self-love a duty but merely ‘permits’
[erlaubt] self-love [Liebe des Wohlwollens] on the condition of one’s ‘being benev-
olent to every other as well’ (MM, 6:451.16–17). Thus the moral permissibility of
self-love as love of benevolence depends on adopting the maxim of practical
love of benevolence towards others (I discuss practical love in more detail in
ch. 4.2.2).

But still, the more virtuous we become, the less we will entertain self-regarding
inclinations, and the more we will be dedicated to pursuing our imperfect
duties of love towards others. Is there some sort of limiting value in our striving
to approximate the highest good as our own moral perfection – a limiting value
where the maxim of practical love towards others has taken over and the maxim
of self-love has been reduced to some absolute minimum, such that it only con-
tinues to exist as a part of the formal condition of our duty to love others?

In the second Critique, the highest good is defined from the perspective of
the agent. It contains the moral law as its determining ground or supreme con-
dition (C2, 5:109) and consists of two elements: virtue (morality) and happiness.
Virtue and happiness have to be combined in the concept of the highest good,
but as they rely on basically opposing principles (those of morality and self-

88 ‘Ich will jedes Anderen Wohlwollen (BENEVOLENTIA) gegen mich; ich soll also auch gegen
den Anderen wohllollend sein. Da aber alle Andere außer mir nicht Alle sein, mithin die Max-
ime nicht die Allgemeinheit eines Gesetzes an sich haben würde, welche doch zur Verpflichtung
notwendig ist: so wird das Pflichtgesetz des Wohlwollens mich als Object desselben im Gebot
der praktischen Vernunft mit begreifen’.
89 ‘unter der Bedingung, daß du auch jedem Anderen wohl willst’.
90 For further discussion of the problem of self-benevolence in this context, see Schönecker
91 There are several aspects of the highest good, which I will discuss in more detail in chs. 3
and 5.
love, respectively), their combination is very difficult to conceive. Kant thinks that if a concept includes two determinations that are necessarily combined in the concept (as must be the case with the concept of the highest good), the determinations must be thought of in terms of ‘ground and consequent’ [Grund und Folge]. According to Kant, this kind of connection is either a logical (analytic) or a causal (synthetic) connection. (C2, 5:111.6–10) Because the two principles in question are heterogeneous, Kant holds that their combination cannot be thought of analytically (in terms of identity) but must be a synthetic relation of cause and effect. This implies that the purely intellectual moral law must be conceived as a cause of sensible happiness. For this to be possible, it must be assumed that one’s moral disposition belongs both to the ‘noumenal’ realm and to the empirical realm of appearances, for happiness as a sensible effect requires that its cause be empirical. (C2, 5:111–115) The type of satisfaction, or the ‘analogue of happiness’ [Analogon der Glückseligkeit] (C2, 5:117.27), that may come about this way is precisely the delight that I previously identified (drawing on the Religion) as negative moral love of delight in oneself. It is the morally acceptable self-contentment felt upon being ‘free from inclinations’ in the sense of being able to override them morally. This type of self-love would therefore be a part of a world in which the highest good, as the subjective combination of virtue and happiness, has been realised. It seems possible, however, that a being that no longer had any inclinations that ran contrary to the moral law would lose its ability to feel satisfaction in being able to override them (but Kant does not address this issue). Recall, moreover, that the evidence I provided for the interpretation that the negative delight of the second Critique is in fact self-love is indirect. In the Religion, Kant both defines moral delight in oneself as self-love and expresses doubts about calling it self-love.

The above considerations to the side, insofar as the highest good involves one’s own happiness and the happiness of others (which it must), it necessarily involves self-love. If we conceive of a world (or better, an infinite approximation towards a world) where our selfish inclinations have been reduced to an absolute minimum, there must still be something that makes us happy. Where can this happiness come from? Since happiness according to Kant’s paradigmatic definition is the sum total of satisfied inclinations, it seems to me that the first main

---

92 Kant calls this the antinomy of practical reason. For a pessimistic view of his discussion, see Beck (1960, pp. 245–247). Here, my intention is not to work with the details of the antinomy as such but only to further articulate the role and structure of self-love in relation to it.

93 The basis of the identification was that the second sense of love of delight in the Religion was Zufriedenheit in oneself based on respect for the moral law, and the second Critique spoke of a Wohlgefallen that was also Zufriedenheit in oneself and also based on respect for the law.
candidate for a possible source of happiness is animal mechanical self-love (enjoyment of nutrition and sexuality, the preservation of offspring, and natural sociality). The second option is that, since humans in such a world would be largely unselfish, a source of happiness might be the benevolent practical love [Liebe des Wohlwollens] that we would shower on and receive in abundance from others around us. It seems possible that there is something intrinsic to loving others as well as ourselves that makes us happy.\textsuperscript{94} In Kant’s philosophy, we are receptive to benevolent love because we wish to be happy or wish to be loved, which is precisely love of benevolence for ourselves. The third option is the moral delight discussed in the above paragraph, which, despite the doubts I have expressed, may plausibly be interpreted in terms of love of delight.

Admittedly, Kant is ambiguous about the nature of the happiness involved in the highest good. Because approximation to the subjective highest good is infinite, it must involve the practical postulate of immortality. This is the only way to conceive of the infinite ascent towards perfect happiness proportionate to perfect virtue in terms of a single agent. But if immortality or the practically postulated afterlife is an immaterial state of being (as we have good reason to assume), and if we know that the ground of inclination is material, then how can the afterlife contain satisfied inclinations? Kant does not seem to consider this problem, especially given the many passages in which there is no implication that the virtuous life is devoid of inclinations.\textsuperscript{95} Our approximation to the highest good may eventually become completely devoid of all inclination-based happiness, but this is unclear; it depends on our interpretation of the postulate of immortality and involves bracketing actual intergenerational moral progress on the species level, where inclinations will clearly continue to exist. Even if we decide to bracket the species level and conceive of approximation toward the highest good as involving the abolition of inclinations, what is left is negative moral delight in oneself, which may arguably be called love of delight for oneself.

\textsuperscript{94} In ch. 4 we will learn that Kant holds that loving others from the incentive of the moral law (with practical love of benevolence) will cause us to love them from inclination too. There is a pleasure connected with beneficent love (LE, 27:419.4–7; cf. MM, 6:402.19–21), and in the highest good as an ultimate end ‘human beings seek something that they can love’ (R, 6:7.27–28) / ‘sucht der Mensch etwas, was er lieben kann’. Neighbourly love from inclination must be subsumed under the basic theorem of self-love (C2, 5:22.6–8).

\textsuperscript{95} However, Kant does problematise his conception in terms of the end of all time. Since the final end of all things is a non-temporal, unchanging state, its possible qualities are beyond our comprehension. In ‘The End of All Things’ [Das Ende aller Dinge], Kant, somewhat jokingly in my opinion, envisions this kind of bliss as the universal congregation’s eternal repetition of the same song (‘Alleluia!’ or monotonous wailing), which indicates a complete absence of change (8:334.36–335.3).
Perhaps the real source of morally deserved happiness lies in all three elements (animal desire, loving others and the need to be loved benevolently, and one’s own virtue). All in all, I find it highly difficult to conceive of the possibility of happiness in the infinite approximation towards the comprehensive highest good without reference to self-love. I call this general finding the persistence of self-love.

1.4 Self-Love Persists on Three Levels

In the first chapter, I argued for a ‘three-level interpretation’ of self-love, according to which self-love can be plausibly divided into three ‘levels’ utilising a vertical metaphor to describe ascending levels of rationality.

I held firstly that there is a low, arational ‘ground level’ of self-love, which I identified as ‘animal mechanical self-love’. Animal mechanical self-love includes the strongest impulses of human nature and divides into love of life (self-preservation), sexual love and parenting (preservation of the species), and sociality. The existence of these terms shows that Kant discusses the strongest impulses of human nature in terms of love. I pointed out how animal mechanical self-love can be understood as desire and how it is generally a predisposition to the good, at least in the sense that it is a necessary condition of species-level moral progress. Further, I showed how the social impulse can be seen as the ground of the unsociably sociable self-love of humanity, which belongs to the history and teleology of moral progress. I also discussed the role of love of honour in this context.

Second, I argued that there is a higher level of self-love that implies reason and which I called the ‘the broad middle level of self-love’. I showed how this level can be analysed in terms of the general division of love into love of benevolence and love of delight. Self-love is fundamentally the principle of one’s own happiness: it is love of benevolence towards oneself, which consists in wanting things to go well for oneself and being loved by others. This love of benevolence is morally permissible on the condition that one includes the happiness of others in one’s maxim. I identified three separate notions of love of delight in oneself: 1) pleasure in love of benevolence towards oneself (in which case love of benevolence and love of delight are equivalent); 2) a feeling of self-contentment (as ‘negative’ freedom over inclinations) or a feeling of positive moral merit due to the performance of imperfect duties; and 3) morally reprehensible arrogance or self-conceit, which is an unjustified claim to respect from others that arises from making the maxim of love of benevolence towards oneself an unconditional practical law.
Third, I argued for the *persistence of self-love* by analysing the place of self-love in what I call the third or highest level of self-love. Unlike the two lower levels of self-love, which are actual, the third level is an idealised notion that includes close proximity to the moral-physical highest good (the virtue and happiness of all rational creatures). Since happiness involves self-love by definition – and since even negative moral self-contentment can plausibly be interpreted as love of delight – I concluded that self-love is an irreducible aspect of even the highest stages of moral happiness. This conclusion is in line with Kant’s view that there is no reason to rid ourselves of inclinations as such. Further, benevolence towards others [*Liebe des Wohlwollens*], which figures as a path towards universal moral happiness, can only be viewed as conducive to happiness under the supposition that human beings want things to go well for themselves or want to be loved by others. In Kant, then, love can make us happy only if an element of self-love persists.