PASSION

The Journal of the European Philosophical Society for the Study of Emotion

Editors in chief
Alfred Archer
Heidi Maibom

Volume: **02**
year: **2024**
Publisher: **Open Press TiU**
ISSN: **2773-1715**

01

Table of content

Locke, Anger, and the State 1
J.K. Numao

Are all emotions social? Embracing a pluralistic understanding of social emotions 15
Gen Eickers

To crave is like to be in love 30
Rachel Frenette

Loving a Narrator 48
Pilar Lopez-Cantero

Received: 13 August 2023
Accepted: 20 October 2023
Published: 30 June 2024

Locke, Anger, and the State

J.K. Numao - Keio University, Japan

Abstract

While "anger" or "angry" are words often used to describe today's societies and politics, they are not the ones that commentators of Locke use when talking about his political theory. This article argues that anger is in fact very much at the heart of his political outlook, and that focusing on this passion helps us to see an important aspect of the state. Since little has been said in relation to Locke and anger, this article starts by examining how he defined, viewed, and justified anger. It shows how Locke conceptualised a problematic component of anger, revenge, in terms of retributive punishment, which, when proportionate, can be viewed positively. The article then shows how Locke placed and tamed anger in the state's hands to resolve the problems in the state of nature caused by anger. In this way, this article sheds light on the state's role as an anger manager.

Keywords: Locke, anger, revenge, nature of the state

1. Introduction

The following is a reflection on the location and role of anger in the political theory of the seventeenth-century English philosopher John Locke. While talk of an "age of anger" is all the rage in the media today, anger has not attracted the same kind of attention in Locke scholarship.[1] Studies that deal more generally with Locke's discussion of passions do exist, often highlighting its Stoic character (Nuovo 2008: 11–12; Kelly 2011: 46–53; Di Biase 2016).[2] Anger is indeed usually mentioned in these studies. However, to my knowledge, there is no comprehensive study of Locke's discussion of anger as of yet. The impression that this lacuna gives is that anger did not occupy the philosopher's mind very much. Against this backdrop, the thesis I wish to put forward, somewhat provocatively, is that anger does figure in Locke's political theory, and moreover that it figures quite centrally. People in the Lockean state of nature experience great inconveniences due to an anger problem: angry people judge and excessively punish others, making those who are punished angry in turn. Locke's solution to

1 For example, "anger" is not listed in the subject index of the *John Locke Bibliography*. See Attig 2023.

2 There are also studies that look at the impact of Locke's ideas on the passions. See, for example, Sodano 2017, which gives an account of the eighteenth-century inheritance of Locke's ideas on the passions. See also Radcliffe 2018: 197–213 for a general survey of the relationship between passions, reason, and action in the seventeenth and eighteenth centuries and Locke's place within this intellectual tradition.

the problem is the creation of the state. Looking at Locke through the lens of anger will thus help us to bring to light an overlooked aspect of the nature of the state as an anger manager.[3]

In the next section, I will precede my discussion of the place and role of anger in Locke's political theory with a more general discussion of how anger is defined, viewed, and justified in his thought.[4] I do this for two reasons. Firstly, as noted, this is because little has been said or done hitherto by way of making a case for a Lockean *account* of anger. I will therefore try to reconstruct and analyse Locke's account of anger, which I believe will be a fruitful and meaningful endeavour in itself. Secondly, and more importantly in connection with my thesis, I begin with a general inquiry into Locke and anger because what anger is and how it plays out in a general context for Locke will help us to see how it fits in his political theory more clearly.

2. Locke on Anger: Definition, Evaluation, and Justification

In this section, I will first briefly introduce Locke's definition of anger. I will then survey what he says about anger. After examining how Locke views anger, I will return to his definition to discuss it more thoroughly, and also look at how he justifies the passion. Locke's most definitive account of anger appears in his philosophical work *An Essay Concerning Human Understanding* (1979. Cited by book, chapter, section), in the context of discussing how different passions are produced, from a reflection on pleasure and pain and their different causes (Radcliffe 2018: 204), Locke offers a series of definitions of various passions, including anger, which he defines as follows: "*Anger* is uneasiness or discomposure of the Mind, upon the receit of any Injury, with a present purpose of Revenge" (1979: 2.20.12). This definition of anger would later be adopted and cited in Samuel Johnson's *A Dictionary of the English Language* of 1755 (Johnson 1755; Dixon 2020: 12).

Martha C. Nussbaum (2016: 5), a modern scholar who has written extensively on anger, holds that there are two components to anger: a) "the idea of a serious wrong done to someone or something of significance," and b) "the idea that it would be a good thing if the wrongdoer suffered some bad consequences somehow." In an article examining the history of the idea of anger, Thomas Dixon (2020: 12) notes that Locke and Nussbaum have "similar" definitions of anger. As can be seen, Locke's definition does indeed include the two elements Nussbaum identifies: a perception of being wronged[5] and the idea of wanting revenge.[6]

Before proceeding to Locke's view of anger, I should note that I will be assuming that, when Locke uses the word "anger" and its cognates, he is using it in the way more or less defined in the manner above. To use Locke's own terminology, I will be assuming that he is not making a rigorous distinction between the technical, "philosophical," and everyday "civil" use of the word (1979: 3.9.3), at least not to an extent that would radically alter the twofold definition of anger.

3 I am indebted to Billy Christmas for his suggestion on how the article's argument might be best framed.

4 I do not discuss the aptness or effectiveness of anger, which are also staples of discussions of anger in the modern literature. See, for example, Srinivasan 2018; Silva 2021.

5 For Locke, injury involves an infringement of one's rights. See Shimokawa 2013: 573.

6 Victor Nuovo (2008: 12) suggests Locke may have borrowed this definition from the Stoics, citing Seneca's "On Anger." See also Dixon (2020: 12); Marshall (1994: 164n). Nussbaum's (2016: 17; 2018: 72–3) definition of anger, by contrast, is modelled on Aristotle's. It might be worth pointing out that Aristotle and Seneca had different views about anger. Whereas Seneca (2010) believed anger should always be avoided, Aristotle (1999: 61 (1125b)) argued that one who "is angry at the right things and toward the right people, and also in the right way, at the right time, and for the right length of time, is praised."

2.1. Locke's View of Anger

Anger has had both its supporters and detractors. Traditionally, however, starting with Seneca's (2010) merciless onslaught, anger has tended to be criticised, and therefore has been on the defence. More recently, though, writers are coming in its defence, neatly captured by a title like *The Case for Rage*, to cite just one example (Cherry 2021). However, for Nussbaum (2016: 5), anger "is always normatively problematic, whether in the personal or in the public realm." This is because the idea of revenge—component b) above—is "normatively problematic," and so with it, anger falls too (Nussbaum 2016: 15).

Given the similarity between Nussbaum and Locke's definitions of anger, it might be supposed that Locke would also have a negative view of the passion. Yet a quick survey of Locke's writing suggests that, while anger on the whole is not good, it is acceptable and called for in some cases. Below, I will compile Locke's various remarks, both negative and positive, on anger from his philosophical, political, educational, and theological writings. After this exercise, I will return to Locke's definition of anger and consider how his way of thinking provides the conceptual tools to justify it.

Let me begin with some of Locke's negative remarks. In the *Essay* itself, anger is considered a "violent passion" that can interfere with our thought (1979: 2.21.53). Likewise, in the *Conduct of the Understanding* (1996, cited by section), passions including anger sometimes "take possession of our minds with a kind of authority," seizing the understanding. These passions are referred to as a form of "tyranny on the understanding," or a "clog" hanging upon the mind, putting people under "the power of an enchantment" (1996: §45). These negative remarks resonate with Stoic scepticism towards passions like anger (Di Biase 2016: 222). Such is the understanding of anger Locke seems to have in mind when he talks about the "chastisement" (i.e. whipping, beating) of an obstinate, crying child in *Some Thoughts Concerning Education* (1989, cited by section). Here, he tells the parents to chastise the child—if it comes to that—"without Passion" and not "in the heat of Anger," which will cloud their judgment and cause them to overdo it (1989: §112). Once children start to see the difference between good and bad behaviour, "Kind concern and help" should be the norm rather than "the Anger and passionate Reproaches of their Tutor and Parents" (1989: §80).

As already noted, Locke's definition of anger includes a revenge component. From this perspective, it seems appropriate that we also consider his views on revenge to clarify his views on anger. As might be expected, Locke does not have a good opinion of revenge either. In the *Essay*, revenge is described as "an impatient desire," which "keeps the *will* steady and intent," and "having once laid hold on the *will*, lets it not go" (1979: 2.21.38). In *Some Thoughts*, the vengeful feelings of both children and parents are reprimanded: "The Accusations of Children one against another, which usually are but the Clamors of Anger and Revenge desiring Aid, should not be favourably received, nor hearken'd to" (1989: §109). In the context of discouraging a random application of corporal punishment, Locke writes that "if you punish a Child so, as if it were only to revenge the past Fault, which has raised your Choler," such punishment would not work (1989: §78).

Revenge is also condemned throughout Locke's political writings. In *A Letter Concerning Toleration* (2010, cited by page number), Locke notes that Christians are "admonished that they abstain from all manner of Revenge (*vindicta*), even after repeated Provocations and multiplied Injuries" (2010: 25). In the *Two Treatises of Government* (1988, cited by treatise, section), like the parents in the example above who punish their child excessively, revenge is brought up in the context of excessive punishment. In the state of nature, where everyone has the right to punish wrongdoers, Locke conjectures that "Ill Nature, Passion and Revenge will carry them too far in punishing others" (1988: II.13). Indeed (and as we shall see later), this is one of the major inconveniences of the state of nature: "Men being partial to themselves, Passion and Revenge is very apt to carry them too far, and with too much heat, in their own Cases" (1988: II.125).

Yet, despite all these negative remarks about anger and revenge, there are passages in Locke's work that seem to condone or approve of anger (I will discuss his views on revenge in Section 2.2). For example, in *Some Thoughts*, Locke advises parents as follows: "Parents should well consider, what Faults in their Children are weighty enough to deserve the Declaration of their Anger" (1989: §60). In the *Reasonableness of Christianity* (1999, cited by page number), commenting on what people who have accepted Jesus as the Messiah were to do, Locke notes that Jesus forbade "not only Murder, but causeless Anger" (1999: 123). What these two passages suggest is that, while some kinds of anger were unacceptable, others could be acceptable. Not all faults called for anger, but some did, being "weighty enough"; similarly, anger that was not "causeless" was not forbidden. Likewise in *A Paraphrase and Notes on the Epistles of St Paul* (1987, cited by page number), Locke renders the well-known passage on anger "Be ye angry and sin not: let not the sun go down upon your wrath" (Ephesians 4:26) as follows: "If you meet with Provocations that move you to Anger, take care that you indulge it not so far, as to make it sinful: Defer not its Cure till Sleep calm the Mind, but endeavour to recover your self forthwith, and bring your self into Temper" (1987: 650).[7] Again, the point is that there is a form of anger that can become sinful, but also, that there could be anger that was not sinful just yet.

Now that we have seen that anger is sometimes acceptable and called for on Locke's account, I want to consider how he could defend this position, especially given that he had a very negative view of one of anger's components, revenge. To do this, I will return to the postponed task of examining Locke's definition of anger.

2.2 Locke's Definition of Anger Considered

To recapitulate, Locke's definition of anger is "uneasiness or discomposure of the Mind, upon the receit of any Injury, with a present purpose of Revenge." I will first say something about the idea of uneasiness in the definition, which will then take us to the idea of revenge.[8] In the *Essay*, Locke tells us that uneasiness is experienced "in the want of an absent good" (1979: 2.21.31, 2.20.6). Uneasiness in turn gives rise to an equal desire to quell it. In fact, Locke says we may call uneasiness "desire," being "scarce distinguishable from it"; all the while there is a desire to ease an uneasiness, we may call it desire, and once this desire is attained, uneasiness is no longer there. If bodily pain is experienced, then ease from pain is the absent good, and there is a "desire equal to the pain" to ease it.[9] Pain, which is uneasiness,[10] is not only about bodily pain (1979: 2.20.15), but also of the "Mind" (1979: 2.20.17, 2.21.31). The uneasiness in anger, according to Locke's definition then, is of the latter kind, being "an uneasiness … of the Mind."[11]

This seems to be the right point at which to bring revenge into the picture. When we think about revenge in the everyday sense as "getting even" with the perpetrator, the language of uneasiness, which uses the language of equalising, makes a great deal of sense. You are wronged and want something that can neutralise the wrong

7 See, for example, Francis Bacon's (1985: 226–7) essay "Of Anger," which cites this verse. On the "delicate distinction between just and sinful anger" dogging Christian writers, see Enenkel and Traninger 2015: 4; Konstan 2020: 108–10).

8 Since this definition of anger is the one Locke uses from the first edition of the *Essay*, it is unclear whether we should make too much of the word "uneasiness," which is introduced in the second edition and is used thereafter to explain the force that determines the will (1979: 2.21.29, 31, 2.20.6). But since he did continue to use the term "uneasiness" in the definition of anger, even after the second edition, we may say that he did not find it incompatible with the discussion of volition in subsequent editions.

9 Locke notes that not all absent good is a pain, in which case desire might not accompany it (1979: 2.21.31). I suppose an example might be something like the following: the presence of a wealthy uncle might be considered a good, but at the same time, his indifference to you might not necessarily be a pain given that you are not destitute, and thus you may not experience any desire to want that uncle in your life.

10 See Locke 1979: 2.20.5. Locke replaces pain with uneasiness in the fourth and fifth edition of the *Essay* when talking about "Uneasiness or Delight." Also see 1979: 2.20.1. Pain is equated with "Trouble," although Locke (perhaps uncharacteristically) does not seem to be too interested in the wording: "call it how you please."

11 Of course, this mental pain may be preceded or accompanied by physical pain.

(Shklar 1990: 94). Could revenge play this role? It could very much for someone like Aristotle. On Aristotle's account (1926: 173 (1378b)), "anger is always accompanied by a certain pleasure, due to the hope of revenge to come." Nussbaum (2016: 17) explains that "the desire for retribution somehow responds to the injury," or more precisely, to the pain produced by the injury. The pleasure of revenge is meant to neutralise the pain. How are things with Locke though? Does revenge have a neutralising effect? To answer this question, we will need to get a better sense of what exactly revenge is for Locke.

In the *Essay*, Locke identifies revenge as a "mixed mode," or a complex idea, made of several kinds of simple ideas (1979: 2.12.5): "To conclude, Let us examine any *Modes of Action*, v.g. *Consideration* and *Assent*, which are Actions of the Mind; *Running* and *Speaking*, which are Actions of the Body; *Revenge* and *Murther*, which are Actions of both together, and we shall find them but so many *Collections of simple Ideas*, which together make up the complex ones signified by those Names" (1979: 2.22.10). As the cited passage shows, revenge involves both a mental and physical component. We have also seen above that revenge is a persistent desire, and causes people to pay back excessively. In the *Two Treatises*, Locke refers to it as an "irregular Passion" (1988: II.199). While these are features of revenge, they do not tell us much about what kind of concept it is exactly. What does the avenger do?

Unfortunately, to my knowledge, Locke does not offer a definition of revenge as such unlike his contemporary Thomas Hobbes. Hobbes (2012: 86), for example, defines revenge (*vindicta*) as follows in the *Leviathan*: "*Desire*, by doing hurt to another, to make him condemn some fact of his own, REVENGEFULNESSE." Also, in his explanation of the seventh law of nature, revenge (*ultio*) is defined as "retribution of Evil for Evil" (2012: 232). The law of nature tells us that revenge should "*look not at the greatnesse of the evill past, but the greatnesse of the good to follow*." Legitimate punishment must be about the "correction of the offender," and without the correction part, that is, "Revenge without respect to the Example, and profit to come," revenge is an act of vain-glory, and ultimately, cruelty (2012: 232). Hobbes's (2012: 542) discussion seems to suggest that revenge that is properly executed can be a form of punishment, although what he says later on seems to come into tension with this statement: "the end of punishing is not revenge [*ultio*], and discharge of choler; but correction, either of the offender, or of others by his example." In saying this Hobbes of course could have meant revenge here as improper revenge, nonetheless leaving room for a legitimate form of revenge.

While Locke lacks a definition along the lines of Hobbes, there are hints in the *Two Treatises* that might help us to get an idea as to what the former could have meant by revenge. Interestingly, what he says in this work seems to suggest that there is a form of revenge that can count as punishment and one that cannot, very much like the reading of Hobbes noted just now. At II.233 in the *Two Treatises*, Locke adduces the absolutist theorist William Barclay to show that even the "great Assertor of the Power and Sacredness of Kings" (1988: II.232) was forced to admit that resistance was permissible in some cases. In discussing Barclay, Locke offers an English translation of the relevant section from Barclay's account of permissible resistance, written in Latin. Locke notes that the "great Advocate of Monarchical Power" attached two limitations to resisting a tyrant: one, you must resist with reverence, and two (and which is relevant for the present case), you must resist "without Retribution, or Punishment" (1988: II.233). In Locke's English translation, while the term "punish" is used once, "retribution" is not used at all. Instead, "revenge" is used three times and appears in the following contexts: 1) self-defence is part of the law of nature, but for the people "*to revenge themselves upon him* [i.e. the King], *must by no means be allowed them*"; 2) the people may repulse the present attempts of the King, but "*must not revenge past violences*"; 3) the people may prevent mischief before it is done, but when it is done, "*they must not revenge it on the King*." "Punish" is used in reference to, or rather, to summarise these acts: "*For it is natural for us to defend Life and Limb, but that an Inferiour should punish a Superiour, is against Nature*."

It is noteworthy that Locke uses the term "retribution" or "punishment" to capture these acts of "revenges" (the use of the terms in the original text – *ultio*, *ulciscendi*, *vindicare* – each standing for "revenge" was not of Locke's own choosing after all). In other words, this move suggests that, in Locke's mind, retribution and revenge contain each other; they may even be interchangeable terms. If so, looking at "retribution" next might shed some light on Locke's understanding of revenge.

Locke uses the term "retribute" twice in the *Two Treatises*, at II.8 and II.176. The first appears in Locke's discussion of the right to punish in the state of nature. Everyone may punish a wrongdoer, but "only to retribute to him" what is proportionate. This passage is cited as an example in the aforementioned English dictionary by Johnson under "retribute," which gives the definition: "To pay back; to make repayment of." Indeed, punishment itself is defined in this section of the *Two Treatises* as "lawfully" doing "harm to another" (1988: II.8). The second use of retribute appears in Locke's discussion of the appeal to heaven. If you appeal to heaven and resist an aggressor, you must be sure that you do have a just cause, otherwise God will "retribute to every one according to the Mischiefs he hath created to his Fellow-Subjects" (1988: II.176).[12] As this will be happening at God's "Tribunal," this retribution of God's may be equated with punishment. Therefore in both cases, on Locke's account, retribution is a form of payback, and is part and parcel of punishment. We may infer from this discussion that, in Locke's thought, revenge is conceptualised as an act of payback which also features in legitimate punishment. It is this possibility that helps him to maintain that revenge, and hence, anger, may be justifiable. The question is, then, what kind of retribution is legitimate?

According to Alex Tuckness (2010: 730), Lockean punishment is legitimate insofar as it is focused on the public good or is used "to protect the rights of citizens." To be legitimate, punishment must be proportionate to the transgression, and what is proportionate is that which serves the end of restraint and reparation (1988: II.8). As for restraint, by punishing the wrongdoer, the goal is to "make him repent the doing of it, and thereby deter him, and by his Example others, from doing the like mischief" (1988: II.8, see also II.12).[13] As for reparation, victims—and only the victims—may demand and recover from the perpetrator the damage that they have caused.[14] Within these bounds, retribution will count as legitimate punishment (See also Locke 1997: 339, "Punitive Justice").

Let us now piece together the preceding discussion. Kept within limits, retribution counts as a legitimate form of punishment. If revenge and retribution contain each other, then just as retribution can be proportionate and legitimate, revenge can exist in a form that is cleansed from its excessive form.[15] Thus, while people can experience anger, and anger includes a revenge component, given that the revenge impulse is tamed, it can be seen as an acceptable passion. Anger can be justified even with the revenge component.

12 In his translation of Locke's *Letter Concerning Toleration*, William Popple renders *rependet* (repay, pay back; reward) in the context of God's judgment as "retribute": "God, I say, is the only Judge in this case, who will retribute unto every one at the last day according to his Deserts" (Locke 2010: 49).

13 It may be interesting to compare Locke with a particular reading of Aristotle (Scheiter 2010: 1, 7, 8). On this view of Aristotle, while punishment and revenge are different, revenge is "analogous" to punishment. Revenge is also about "preventing future injustice," not dissimilar to what Locke says. Moreover, for Aristotle, revenge need not be severe; what is needed is pain that can cause people to repent. Yet the difference between revenge and punishment for Aristotle is that revenge is about "preventing injustice from happening to *us*," that is to say, the person seeking revenge.

14 Indeed, John Dunn (2003: 60, 62) refers to the right of reparation as "the right of individual *revenge*" or "the right to *avenge* individual injuries" (emphasis added).

15 Thinkers today, such as Robert Nozick (1981: 368), argue that revenge and retribution are distinct. Nozick does admit that retribution and revenge "share a common structure," namely "a penalty is inflicted for a reason . . . with the desire that the other person know why this is occurring and know that he was intended to know." Yet Nozick maintains they are not the same. See also Walen 2020. For Locke, revenge and retribution are concepts that contain each other, and so we may say that he does not share Nozick's view. Also, against Nozick's view, see Fassin 2018: 36.

We may add here that the idea of revenge as a form of retributive punishment suggests that it is for the most part meant to be applied to inter-personal relationships; or, to put it another way, it is meant to hold *someone* rather than *something* morally accountable (cf. Callard 2020: 18). For example, tennis players sometimes throw their racket in anger, blaming it for their bad performance, thereby seeking revenge on it metaphorically. If, however, revenge is about punishment, then on Locke's account it will not make sense to say that you are punishing the racket. Just punishment, for Locke, implies a moral agent, someone who could have done something otherwise but did not, while the racket in this specific case could not have of its own power (1979: 4.17.4).[16] Locke did, of course, entertain the possibility that there could be rational monkeys (1979: 3.11.16), and so technically speaking, punishment may, in theory, be justly inflicted on non-human animals or things. Perhaps a case can be made for this with respect to, say, AI robots in the future, but it seems to make sense to focus on revenge as applying to other human beings for the most part in the present discussion. Seeing revenge, and hence anger, in this light will help us to appreciate it in the next section as a political concept, and the appropriateness of its mobilisation and use by the state.

Before ending this section, I want to rewind a little bit and go back to the discussion about the relationship between uneasiness and revenge, now that we have a better sense of Lockean revenge. The question under consideration was: what is the relationship between revenge and the pain one experienced, accompanied by the desire to quell it? Speaking of (envy and) anger, Locke argues that they are not caused by "Pain and Pleasure simply in themselves, but having in them some mixed Considerations of our selves and others, are not therefore to be found in all Men, because those other parts of valuing their Merits, or intending Revenge, is wanting in them" (1979: 2.20.14). Given that Locke says other passions, such as love and hate, that terminate "purely in Pain and Pleasure" are "found in all Men," it seems he is saying that the intent of revenge is distinct from mere pain, and on that account anger is not experienced by every person.[17]

There may be a number of ways to interpret Locke here. First, he could be saying that anger is about wanting a wrong to be righted *and also* wanting revenge *separately*. While the first part might be experienced by all people, the second might not. Thus, the desire for relief from the injury, or pain, and the desire to exact revenge, are different. For this reason, anger is not universally experienced. If this is the case, then the equalising "getting even" metaphor seems to collapse. There is a desire for something to equal the wrong (and to that extent, the idea of equality might be there), but this is not through revenge. Revenge on this account would be a random, independent act.

In a similar vein to this possibility is the following. Anger is a sense of being wronged, accompanied or expressed by a desire for revenge. In his discussion of the parents' declaration of anger in *Some Thoughts*, Locke rephrases anger as the parents' "Displeasure" being declared "to a Degree that carries any Punishment with it." He also talks about the "Marks" of displeasure (1989: §60). Perhaps there is displeasure accompanied by a displaying of one's eagerness to exact revenge, which amounts to anger, and simple displeasure, which does not entail this expression of vengefulness. Not everyone might express their displeasure to such a degree, and as a result might not experience anger. Revenge would be less of a random act being a mode of expression of anger, but, insofar as the equalising effect of uneasiness is concerned, it is a separate act. Nothing is being satisfied through it.

16 It is possible to see this in the following way too. You could say that you are seeking revenge on yourself by destroying your property, i.e., your tennis racket.

17 Di Biase (2016: 221) refers to the "special status" of anger. Lehmann (2015: 18, 22) suggests that anger is not universally experienced because, for Locke, anger "counts among the privileges of an honourable man: without a desire for revenge there can be no anger," where revenge is for the injury done to one's socially established rights. And alas, not everyone can enjoy these privileges. On this account, Locke comes close to Aristotle's outlook, where the "powerless are incapable of feeling anger" (Konstan 2020: 103).

The third possibility is precisely one in which revenge satisfies something. Like Aristotle, Locke could be saying that anger is about wanting a wrong to be righted and wanting to do that *through* revenge. But because not everyone wants to right a wrong through revenge, anger is not experienced by everyone. The "getting even" metaphor fits the language of uneasiness and equalness. This reading suggests that the intent of revenge and pain are intertwined. Also, this reading makes sense if anger is a "mixed mode," a complex idea made of several kinds of simple ideas (1979: 2.12.5). It might be pointed out, against this understanding of anger being a mixed mode, that Locke says that pain is a simple idea, and that passions are "Modifications" thereof, and moreover that his discussion of mixed modes is introduced only later in II.xxii (1979: 2.20.5). Yet Locke also does seem to suggest that anger is not merely a variation of pain. In this case it would be a simple mode, but mixes another element—revenge—with it. As we have seen, revenge is itself a mixed mode. If this is the case, pain and revenge can both be parts of the complex idea of anger.

This third possibility makes greater sense given the nature and function of Lockean revenge. Revenge as retribution is about inflicting harm on others so as to prevent them and others from harming you and others in the future. It is also about recompensing the harm done to you. In these ways, it does give you pleasure in that you can expect a more peaceful and safer future, and also in that you can regain your former possessions and capabilities, albeit by other means. In this proper form of revenge, you are not meant to get pleasure in the Aristotelian or sadistic sense by getting to inflict pain on others to satisfy your bloodthirstiness or simply for the fun of it (Nozick 1974: 138; Simmons 1992: 157). While the sadist's revenge may be proportionate—namely, the perpetrator may have deserved the degree of punishment inflicted by the sadist—motivation also matters for an act to count as punishment (Simmons 1992: 157). Of course, things may not always be perfect, and people's vengeful impulses may exceed the proportionality needed to quell the original uneasiness caused by the pain, or may derive pleasure for the wrong reasons. Nevertheless, when our vengeful appetite operates correctly, the language of uneasiness captures what is going on when we are angry. The question then is how we can vent our anger properly. This brings us to Locke's discussion of the state.

3. The Anger Problem in the State of Nature and the State as an Anger Manager

The story of the origin of the state, for Locke, begins in a pre-political state of nature. It is a state of perfect freedom and equality of all people under the law of nature (1988: II.4). The state of nature is not the state of war for Locke as it was for Hobbes, but it is still a state of "inconveniences" (1988: II.13, 19, 127). Locke gives three reasons for why people find themselves in this "ill condition." The first reason is that there is no set of laws that people have agreed upon to be the "Standard of Right and Wrong" (1988: II.124). True, the law of nature exists, and people can in principle know what it demands, but people often either fail to apply the law to themselves or, more fundamentally, fail to make an effort to learn what it requires of us. The second reason is that there is no impartial "Judge and Executioner of the Law of Nature" (1988: II.125). While everyone has the right to play this role in the state of nature, "Passion and Revenge" makes people overdo punishment, particularly when a case involves themselves. Lastly, sometimes people will not be able to execute punishments simply due to a lack of power (1988: II.126).

We see that even if people know what the law of nature demands, and have the means to punish the perpetrator, being who they are, emotional, they could show favour towards themselves and their close ones and tend to be harsher to others, this harshness precipitated by vengeful passions such as anger (1988: II.13). There is a problem in the state of nature caused in part by anger and its effects. Even if two key obstacles can be overcome, there is

the third (by which I mean the second reason above) that renders life in a state of nature inconvenient—the anger problem, wherein angry people judge others, causing anger in those who are subject to this judgment.

The solution Locke proposes is the state. Let us see then what the state can do to remedy the anger problem. In political society, the legislative will create standing laws, and indifferent judges will judge according to these laws (1988: II.131). Wronged and angry people will not themselves be judging their own cases. When wronged, people will have a place to make an appeal. That said, there are cases even in political society where you may lack this relief. The particular case Locke has in mind is when a thief is robbing you. In this case you lack political and legal relief because you do not have time to appeal to the common judge (1988: II.19, 207). But on the whole, you can and do enjoy the benefit of appealing to the law in political society, and moreover you are expected to do so rather than taking matters into your own hand, as the right to punish is given up to the body that is appointed by the members (1988: II.127). Members must now use their natural force "to assist the Executive Power of the Society" (1988: II.130). The state is equipped by angry people, and on their behalf exercises anger and revenge against wrongdoers at a proportionate level.

In a sense, therefore, a Lockean state is not unlike an Aeschylean one, or at least on one reading of Aeschylus. To illustrate this point, let me compare Locke's position to an example that Nussbaum often cites, the ending of Aeschylus's *Eumenides*, the third play in the *Oresteia*.[18] In this tragedy, Orestes, the son of Agamemnon, is being pursued by the Furies, the ancient goddesses of revenge, for killing his mother, Clytemnestra, and her lover, Aegisthus, who were responsible for murdering Agamemnon, Clytemnestra's husband and Orestes's father. Through the intervention of the god and goddess Apollo and Athena, Orestes is set to be judged by the citizens of Athens in court rather than falling into the vindictive hands of the Furies. The jury gives its verdict, and the votes are split. However, Athena casts her vote in favour of Orestes, and he is acquitted. The Furies are angry. Yet Athena appeases the Furies by offering them a venerable place in the city.

According to the traditional interpretation, the "Furies agree to accept the constraints of law, but they retain an unchanged nature, dark and vindictive". What this means is that the Furies must be honoured as they are (Nussbaum 2018: 65). Nussbaum disagrees with this interpretation. Nussbaum notes that, at the end of play, the Furies are not only received into the city; they are also transformed. Indeed, it is this transformation from the Furies to the "Kindly Ones" that is important and underlies what she calls "Transition-Anger." The message, according to Nussbaum (2018: 67), is that while the Furies (i.e. revenge) "are still needed" given an imperfect world, "they are not wanted or needed in their original form."

My reading is closer to the traditional reading, and this helps us to see the nature and function of the Lockean state. The first point to highlight is that the Furies are invited and do accept a place in the city, leaving aside whether their nature is changed or not. The Furies, moreover, will be in the city eternally. The reception and inclusion of the Furies suggests that the revenge component is not incidental, but is an integral part of the city's order (Aeschylus 1979: lines 891, 897). People's impulses to get back at the wrongdoer must be taken seriously and cannot be brushed aside easily. Whether outside or inside the courtroom, people need and will seek some sort of payback. Likewise, Locke retains the revenge component in his definition of anger as a way to hold people accountable; moreover, he equips the state with the power to mobilise anger against wrongdoers, or more simply, to retribute them.[19]

18 I want to emphasize that I am bringing this play up only to illustrate Locke's position clearer. My concern is not to offer an accurate interpretation of the play. I neither pretend to do so, nor do I believe I have to in order to make my point.

19 Again, it is interesting to compare Locke's position to Hobbes's. Hobbes (2012: 84) defines anger as "Sudden *Courage*." Courage, which is also defined, is fear "with hope of avoyding that Hurt by resistence." See also Hobbes's (1966) discussion in *The Whole Art*

On Nussbaum's account, the Furies, however, are transformed, and do not remain bloodthirsty and vindictive. They embrace love. Similarly in Locke's case, the revenge component is weakened and the bloodthirstiness is stripped away at the political level. However, this is where I part ways with Nussbaum's reading. Yes, the Furies make some change, but they are neither completely transformed into agents of love as Transition-Anger might suggest, nor do they lose their previous *lex talionis* mentality (Nussbaum 2018: 94). The Kindly Ones' words for the citizens, "But may each give joy to each" (Aeschylus 1979: 219, line 984), a passage Nussbaum cites, is rendered "may they return joy for joy" (Aeschylus 1971: 365) or "may they reciprocate joys" (Aeschylus 2002: 111) in different editions.[20] Also, Athena remarks, right after this benediction, "From these fearsome faces / I see great good for these the citizens. / For if, kind in return for kindness, you do them ever great honor" (Aeschylus 1979: 219, line 993). What this remark indicates is that, in return for the Kindly Ones' kindness, if people honour them, the city will prosper: the Kindly Ones will in turn show the people kindness. These passages together suggest that the mentality of *lex talionis*, returning good with a proportionate good or bad with proportionate bad, is very much alive in the city. Furthermore, the Kindly Ones reciprocate, but they are not all about reciprocating love: they say: "May they [i.e. citizens] hate with one accord" (Aeschylus 1979: 219, line 986). Thus, hating, or returning enmity with enmity, is still on the list of activities.[21] What is important is that we hate with one heart.[22]

Applying this to Locke's case, revenge is there, but altered. However, it need not be transformed completely into love. The revenge that can constitute a justified anger is one that is not bloodthirsty but is proportionate to the wrong experienced, and that can secure the citizens' rights by preventing future wrongdoings and exacting reparation from the perpetrators. Inflicting harm, or wanting to do so, is appropriate insofar as it is within these bounds.

Therefore, on Locke's account, anger is something to be tamed, not eradicated. This perspective allows us to see why anger can be both the problem and solution in Locke's political theory. On the personal level in the state of nature, there may be both acceptable and unacceptable anger. However, people in this state also have executive power. When this power is combined with bad anger, then it creates the anger problem—angry people judging their own cases, fuelling further anger in society. Thus, anger needs to be tamed, and this is achieved through the creation of the state. The state places the executive power in the hands of impartial judges. The point of the state is to try to put the executive power in the hands of a third party who do not have a conflict of interest. While there may be people who experience anger, they will not be the ones who will be judging their own cases.

In this respect, anger seems to be taken away. Nevertheless, it can be said that those who exercise the executive power are engaged in a form of anger. They observe a wrongdoing, hold someone accountable, and give

of Rhetoric, an abridged translation of Aristotle's *Rhetoric*. So, while anger includes a resistance component, it does not include a revenge component (Jaede 2017: 841). For Hobbes, anger is a "primordial passion that aims at overcoming any kind of hindrance, including enemies threatening to inflict hurt" (Jaede 2017: 841). By doing so, he "distinguishes anger from revengefulness" (Jaede 2017: 842). By contrast, revenge according to Hobbes (2012: 112), or rather, "excessive desire of Revenge" is "Rage," which is "Madnesse." Revenge is associated here with anger, but anger that is out of control. Thus, while anger itself is a passion to fight off an obstacle, combined with revenge, it becomes an unjustifiable madness. Anger must be purified of revenge to be viable.

20 The term is *antididoien* (ἀντιδιδοιεν), which may be rendered in any of the ways cited here.

21 Writers often distinguish between hatred and anger, as does Nussbaum. But she also notes how, when we focus on revenge, or the payback impulse, "things are more complicated," for "wanting payback looks like a kind of hatred of the person" (Nussbaum 2016: 50). For a more sceptical view, see Dixon (2020: 13–16).

22 Harris (2001: 162) suggests that the end of the play is about how anger should not be directed at each other but rather how the Athenians "are to concentrate on foreign wars (which recalls the beginning of the trilogy)," i.e., the Trojan War. The Trojan War itself was an expedition "commanded by Zeus to avenge the violation of a crime against hospitality" (Nussbaum 2001: 33).

condign punishment. Anger is not removed from the state, but rather managed. Since the punishment it delivers is controlled, the state puts an end to the vicious cycle of excess punishment leading to further acts of retaliation. The state manages the anger exercised by the executive power, and through this minimises the anger-provoking moments that previously existed in the state of nature. Moreover, through the creation of the state, people have a place to take their anger. The courts may be a place full of anger, but cases will not be judged by people who suffer from the anger problem.

It should be noted that the creation of the state in and of itself will not entirely do away with the anger problem. Tyrants may rule, and, when affected by anger, may behave worse than people in the state of nature, thus making the situation worse than a merely inconvenient state (1988: II.199, 225). Unless an appropriate ruler is in power, the anger problem will remain, even under a government. The solution in this case is to overturn the government. However, Lockean citizens now seem to have a different problem: they tend to be too patient and will not rise up easily (1988: II.224-228). The goal of overturning the unjust government, however, is not to return to the state of nature, but to replace the unjust ruler with an appropriate one. While the state itself is susceptible to the anger problem, this susceptibility does not render it useless. It is still the solution. A ruler is needed—just not a tyrant.

4. Conclusion

Anger has been a neglected topic in Locke studies. Therefore, one aim of this article has been to fill in this gap. As we have seen, Locke's general view of anger was that it is for the most part not a good thing, but that in certain cases it is justifiable. The justifiability of anger in turn seems to hinge on one of the components of anger, the revenge component. Proportionate and controlled revenge can be a form of legitimate retributive punishment, and insofar as the desire for revenge remains within this scope, anger is justifiable.

Channelled through the state, our vengeful appetite is tamed. Just as it is appropriate to get angry at injustice and seek revenge proportionately, so the state will be expected to undertake this task. While, in the state of nature, things could get too personal and out of hand, the state is at an advantage because it can adopt an impartial perspective. The state is itself a product of the vindictive power of the people in the state of nature, but one that is regulated. The state uses anger's energy and makes it its solution to the anger problem. There is still the possibility that the executive may suffer from uncontrolled anger, and so the people need to remain vigilant.

The above consideration has also shed light on an oft-unsung aspect of the function of the state, namely as an anger manager. Locke states in the *Two Treatises* that the reason why governments are formed is to enable people to preserve their property, property being defined broadly as their 'Lives, Liberties and Estates' (1988: II.123-124). The state has the power to enact laws and punish those who break them. The state also protects people from foreign injury (1988: II.3). Quite naturally, common descriptions of the Lockean state have tended to focus on its character and identity as an "umpire" or "judge" (Grant 1987: 74; Ashcraft 1987: 8; 1994: 248), "common executor of the law of nature" (Mack 2013: 77), or a "protective agency" (Nozick 1974: 22).[23] Focusing on the state's function as an anger manager does not prompt us to deny any of these characteristics. The Lockean state does serve as a judge; it does execute the law of nature; it does protect people's rights. While confirming these roles, the article's

23 For how these different roles might be connected, see Simmons 1992: 313–17.

focus expands our understanding by showing us more specifically and fundamentally why and how the state goes about doing these things—by identifying anger as a crucial problem and by managing it.[24]

As noted, the article shows that a key role of the state is to ensure impartiality of the power exercised over others so that those who are subject to that power will have no good cause to retaliate in return. Importantly, while the state may be impartial, it is not entirely impersonal. Qua state, it experiences anger and releases that anger. To that extent, it is emotional. However, it is meant to control that anger in order to control the overall level of anger in society, thereby maintaining peace and order. All this serves as a good reminder that governments are run by people who have ups and downs. Yet, by creating the right conditions, people can use their emotions and natural powers in the right way. Hereby, government by *people*, *qua* emotional being, for *people* likewise emotional is made possible.

Acknowledgments

I would like to thank everyone who participated in the Locke and Lockeanism Working Group Annual Meeting on May 18th and 19th, 2023 at New York University School of Law, Classical Liberal Institute, for their invaluable feedback on an earlier version of this article. I would especially like to thank Billy Christmas, Brian Smith, Daniel Layman, and Douglas Casson, whose suggestions and comments helped me reframe and revise this paper. I would also like to thank Alfred Archer and two anonymous referees for their comments. Finally, this research was supported by Keio Gijuku Academic Development Funds.

24 More generally, studies tend simply to note that the state exists to protect property or that it exists for the public good (e.g. Tully 1993: 14; Harris 1998: 241–2). They may go further to say to what end the state does this, namely to help achieve God's intention to preserve humankind (Dunn 1969: 124–5; Harris 1998: 241). Again, I am not contesting the validity of these points. These studies, however, tend to pass over the role of anger in creating a situation in which people's property become precarious in the first place (Harris 1998: 242). Dunn (1969: 126) does talk about how people's sinfulness can affect human relationships in the state of nature, but the problem is nevertheless not pinned on anger (See also Hampsher-Monk 1992: 102-03). And if anger is not identified as a problem, then it is no surprise that managing it is not noted as the solution, and hence as one of the functions of the state.

References

Aeschylus.1971. *Aeschylus*. Vol. 2. Translated by H. W. Smyth. Harvard University Press.

———. 1979. *The Oresteia*. Translated by H. Lloyd-Jones. Bloomsbury.

———. 2002. *Oresteia*. Translated by C. Collard. Oxford University Press.

Aristotle. 1926. *The 'Art' of Rhetoric*. Translated by John H. Freese. Harvard University Press.

———. 1999. *Nicomachean Ethics*. Translated by T. Irwin. Hackett.

Ashcraft, R. 1987. *Locke's Two Treatises of Government*. Allen and Unwin.

———. 1994. "Locke's Political Philosophy." In *The Cambridge Companion to Locke*, edited by V. Chappell, 226–251. Cambridge University Press.

Attig, J. C. 2023. "John Locke Bibliography." *PSU*, October 10, 2023. https://openpublishing.psu.edu/locke/bib/index.html.

Bacon, F. 1985. *The Essays*. Edited by J. Pitcher. Penguin Books. Original work published 1625.

Callard, A. 2020. *On Anger*. Boston Review.

Cherry, M. 2021. *The Case for Rage*. Oxford University Press.

Enenkel, K. A. E. and A. Traninger. 2015. "Introduction: Discourses of Anger in the Early Modern Period." In *Discourses of Anger in the Early Modern Period*, edited by K. A. E. Enenkel and A. Traninger, 1–15. Brill.

Fassin, D. 2018. *Will to Punish*. Edited by C. Kutz. Oxford University Press.

Grant, R. W. 1987. *John Locke's Liberalism*. The University Press of Chicago.

Di Biase, G. 2016. "John Locke's Stoicism: Grief, Apathy and Sympathy." *Ethics & Politics* 18 (2): 215–39.

Dixon, T. 2020. "What is the History of Anger a History Of?" *Emotions: History, Culture, Society* 4: 1–34.

Dunn, J. 1969. *The Political Thought of John Locke*. Cambridge University Press.

———. 2003. *Locke: A Very Short Introduction*. Oxford University Press.

Hampsher-Monk, I. 1992. *A History of Modern Political Thought*. Blackwell.

Harris, I. 1998. *The Mind of John Locke*. Rev. ed. Cambridge University Press.

Harris, W. V. 2001. *Restraining Rage*. Harvard University Press.

Hobbes, T. 1966. "The Whole Art of Rhetoric." In *The English Works of Thomas Hobbes*, vol.6, edited by W. Molesworth, 419–510. Scientia Verlag Aalen. First published 1637.

———. 2012. *Leviathan*. Vol. 2. Edited by Noel Malcom. Oxford University Press. First published 1651.

Jaede, M. 2017. "Forgiveness and Reconciliation in Hobbes's Natural Law Theory." *History of European Ideas* 43 (8): 831–42.

Johnson, S. 1755. *A Dictionary of the English Language*. s.v. "anger, n.s.1755." Accessed on March 8, 2023. https://johnsonsdictionaryonline.com/1755/anger_ns.

Konstan, D. 2020. "A History of Anger." In Callard 2020, 101–11.

Lehmann, J. F. 2015. "Feeling Rage: The Transformation of the Concept of Anger in Eighteenth Century Germany." In *Discourses of Anger in the Early Modern Period*, edited by K. A. E. Enenkel and A. Traninger, 16–45. Brill.

Kelly, D. 2011. *The Propriety of Liberty*. Princeton University Press.

Locke, J. 1979. *An Essay Concerning Human Understanding*. Edited by P. H. Nidditch. Clarendon Press. First published 1690, *recte* 1689.

———. 1987. *A Paraphrase and Notes on the Epistles of St Paul*. 2 vols. Edited by Arthur W. Wainwright. Clarendon Press. First published 1705–1707.

———. 1988. *Two Treatises of Government*. Edited by P. Laslett. Cambridge University Press. First published 1690, *recte* 1689.

———. 1989. *Some Thoughts Concerning Education*. Edited by J. W. Yolton and J. S. Yolton. Clarendon Press. First published 1693.

———. 1996. *Some Thoughts Concerning Education and Of the Conduct of the Understanding*. Edited by R. W. Grant and N. Tarcov. Hackett. First published 1697.

———. 1999. *The Reasonableness of Christianity*. Edited by J. C. Higgins-Biddle. Clarendon Press. First published 1695.

———. 2010. *A Letter Concerning Toleration and Other Writings*. Edited by M. Goldie. Liberty Fund. First published 1689.

Mack, E. 2013. *John Locke*. Bloomsbury.

Marshall, J. 1994. *John Locke: Resistance, Religion and Responsibility*. Cambridge University Press.

Nozick, R. 1974. *Anarchy, State, and Utopia*. Basic Books.

———. 1981. *Philosophical Explanations*. The Belknap Press.

Nuovo, V. 2008. "Aspects of Stoicism in Locke's Philosophy." In *Studies on Locke: Sources, Contemporaries, and Legacy*, edited by S. Hutton and P. Schuurman, 1–25. Springer.

Nussbaum, M. C. 2001. *The Fragility of Goodness*. Cambridge University Press.

———. 2016. *Anger and Forgiveness*. Oxford University Press.

———. 2018. *The Monarchy of Fear*. Oxford University Press.

Radcliffe, E. S. 2018. *Hume, Passion, and Action*. Oxford University Press.

Scheiter, K. M. 2010. "Aristotle on the Purpose of Revenge." In *Best Served Cold: Studies on Revenge*, edited by S. C. Bibb and D. E. Montiel, 3–12. Inter-Disciplinary Press.

Seneca, L. A. 2010. *Anger, Mercy, Revenge*. Translated by R. A. Kaster and M. C. Nussbaum. The University of Chicago Press.

Shimokawa, K. 2013. "The Origin and Development of Property: Conventionalism, Unilateralism, and Colonialism." In *The Oxford Handbook of British Philosophy in the Seventeenth Century*, edited by P. Anstey, 563–86. Oxford University Press.

Shklar, J. N. 1990. *The Faces of Injustice*. Yale University Press.

Silva, L. 2021. "The Efficacy of Anger: Recognition and Retribution." In *The Politics of Emotional Shockwaves*, edited by A. Falcato and S. Graça da Silva, 27–55. Palgrave Macmillan.

Simmons, A. J. 1992. *The Lockean Theory of Rights*. Princeton University Press.

Sodano, J. P. 2017. "Uneasy Passions: The Spectator's Divergent Interpretations of Locke's Theory of Emotion." *The Eighteenth Century* 58 (4): 449–67.

Srinivasan, A. 2018. "The Aptness of Anger." *Journal of Political Philosophy* 26 (2): 123–44.

Tuckness, A. 2010. "Retribution and Restitution in Locke's Theory of Punishment." *Journal of Politics* 72 (3): 720–32.

Tully, J. 1993. *An Approach to Political Philosophy*. Cambridge University Press.

Walen, A. 2020. "Retributive Justice." In *The Stanford Encyclopedia of Philosophy*. Edited by E. N. Zalta and U. Nodelman. Winter 2023 edition. https://plato.stanford.edu/archives/win2023/entries/justice-retributive/.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Received: 23 February 2023
Accepted: 11 April 2023
Published: 30 June 2024

Are all emotions social? Embracing a pluralistic understanding of social emotions

Gen Eickers - University of Bayreuth, Department of Philosophy, Bayreuth, Germany

Abstract

While the importance of social emotions is widely recognised, the question of whether all emotions are social, and what this would mean for the category "social emotions," has yet to be addressed systematically. Emotion theorists and researchers have so far proposed different candidates for social emotions. These include non-basic emotions, self-conscious emotions, higher-cognitive emotions, and defining social emotions via their social functions. This paper looks at these different candidates for social emotions and briefly discusses their issues. Discussing the candidates and their issues motivates embracing a pluralistic approach to social emotions. In a further step, this paper will look at approaches which explore how social factors impact emotions in general. This will serve as a basis for explaining what it would mean to propose that all emotions are social. After reviewing different candidates for social emotions, and looking at the impact of social factors on emotions in general, the paper will propose that a pluralistic understanding of social emotions is needed to embrace the different ways in which social emotions may function. Embracing a pluralistic understanding of social emotions does not mean, however, that all emotions are social emotions.

Keywords: social emotions, emotion theory, social norms, social construction, basic emotions, non-basic emotions, social function of emotion

1. Introduction

In emotion theory, there is much talk of "social emotions," but researchers have not found common ground on what this term means. Major debates in emotion research revolve around the question of how culture and sociality influence how we feel, and the ways we express our emotions, but the concept of social emotions often goes unanalysed. When definitions are available or inferable, disagreement abounds. The value and import of the term "social emotions" deserves attention. I will ask: is there a specific subcategory called social emotions that can be clearly distinguished from non-social emotions? Surprisingly, there may be no satisfying way to carve out a special class. I will briefly look at different candidates for social emotions in order to motivate a pluralistic account of social emotions. I will present some issues that the accounts face, but will not argue for or against either of them. Rather, after providing an overview of different candidates for social emotions, I

will provide reasons why a pluralistic account of social emotions should be embraced. A pluralistic account of social emotions may consist of any combination of the criteria offered by the candidates reviewed in Section 2.

As its title indicates, this paper goes further in asking whether all emotions are social. This idea has been pushed by views that emphasise social constructionist understandings of phenomena, particularly of emotions (see Leys 2017). I will draw on social constructionist accounts and argue that recent research developments invite the idea that all emotions are social in a variety of different ways, which places further emphasis on embracing a pluralistic account of social emotions.[1] However, I will also pose the question of whether the fact that all emotions are social in some way means that all emotions are social emotions.

But why is it important to elaborate on social emotions since there are general accounts of emotions already? One could argue that social emotions are implicitly addressed by talking about emotions in general, and that general accounts of emotion have (more or less) implicitly acknowledged the inherent sociality of at least some emotions. But implicitly addressing social emotions and the sociality of emotions does not provide an answer to the question of whether all emotions are social. This question deserves to be addressed, however, since an explicit analysis thereof will help move forward the debate around what social emotions are. It will also address the nature versus nurture debate that is frequently present in general emotion theory and philosophy of science.

In the following, I will (1) look at how social emotions have been categorised in the emotion literature, (2) analyse how emotions are impacted by the social, and (3) will thereby present reasons to embrace a pluralistic account of social emotions, albeit one that remains sceptical about classifying all emotions as "social emotions."

2. Candidates for Social Emotions

What counts as a useful basis for distinguishing between social and non-social emotions? Emotion theorists and researchers, as well as theorists working on issues other than emotions, differ in what they conceive of as "social." There are views and disciplines that purport that everything is social (think of sociology, for example). But there are also views that try to limit the social (think of views defending biological determinism, for example). Views along these lines have been existent in emotion theory as well (for a discussion, see e.g. Leys 2017). So, an emotion might be about something social, affected by social factors, communicated to other people, etc. But does the involvement of social factors in any of these different ways equally make an emotion a social emotion? Furthermore, it seems obvious that at least some cases of even supposedly non-social emotions (e.g., disgust, fear) are caused by or directed at social factors, but does that mean that disgust and fear are social emotions after all?

In the following, I will briefly look at different possible candidates and criteria for the category of social emotions, all of which have been referred to in debates about social emotions or the sociality of emotions. All of the candidates reviewed in the following are controversial, meaning, whether they sufficiently conceptualise the category "social emotions" is up for debate. Let's look at the different candidates for social emotions in detail to figure out whether there is a need for a pluralistic understanding of social emotions.

1 An issue not discussed in this paper is whether the emotion types distinguished here (anger, fear, etc.) themselves count as cultural and normative constructions (see, e.g., Russell 2003).

2.1 Non-Basic Emotions

One way to implicitly tackle the question of what social emotions are has been to portray a dichotomy between basic emotions and non-basic emotions. One can then say that social emotions are non-basic emotions. But what are basic and non-basic emotions?

2.1.1 Origins of the Basic Emotion Debate: Ekman and Darwin

In psychology and neuroscience, a prominent approach to basic emotions is the Affect Program Theory (Ekman 1972; Izard 1977). According to Affect Program Theory, basic emotions are distinctive, biologically based, universal signals: i.e., basic emotions are accompanied by expressions that serve to inform others about the felt emotion; basic emotions also feature distinctive patterns of autonomous nervous system activity; and basic emotions have a quick onset and are of brief duration (Ekman 1999: 52; Ekman and Cordaro 2011). The six emotions that have been most intensively studied by Ekman in this framework are: disgust, anger, fear, joy, sadness, and surprise (Ekman 1972; though see Ekman 1999 for a longer list). According to Ekman, basic emotions are universal. This universality claim is largely based on Charles Darwin's preceding work (1872). According to the standard reading of Darwin's view, each emotion elicits an emotional expression, and some expressions are universal—which in turn points to the universality of the underlying emotions. Ekman's main conclusion, following Darwin, is that we have six fixed basic emotions that are universal, i.e., shared across cultures. Ekman not only follows Darwin, but also conducted cross-cultural studies in order to examine the universality of facial expressions (Ekman and Friesen 1971; Ekman 1972). Ekman tested recognition of the six basic emotional expressions among the Fore, a group in Papua New Guinea. Ekman's main claim has been that the Fore assigned the same meanings to the basic emotion faces as individuals in a Western comparison group. This is the most relevant empirical evidence that he has presented in favour of his account. However, his findings have been criticised, and his account is considered outdated by many scholars.[2]

2.1.2 Non-Basic Emotions as Social Emotions?

There has been a lot of controversy surrounding the distinction between basic and non-basic emotions, and the debate is messy since, often, terms like "non-basic emotions," "higher-cognitive emotions," and "self-conscious emotions" are used interchangeably.[3] Theorists who work with the concept of basic emotions (Ekman 1999; Izard 1977; 2011; Panksepp 2007) have yet to arrive at a common ground regarding which emotions to consider as basic and which emotions to consider as non-basic. Indeed, Ekman thinks all emotions are basic—that is, every genuine emotion will turn out to be innate (Ekman 1994; 1999). Among theories assuming there are both basic and non-basic emotions, there is no consensus over which emotions are non-basic, and so there is no consensus over which emotions are social.

Still, there is a view at least implicit in some of this literature that basic emotions are non-social and non-basic emotions are social. Non-basic emotions have often been characterised as developing later in ontogeny than basic emotions (see e.g. Buck 1999) and to be more complex than basic emotions (Barrett and Campos 1987). Buck (1999), for example, considers the bio-physical structures of emotions to be primary, and argues that social emotions developed from earlier and more basic emotions. Some accounts declare that basic emotions are naturally given, and that non-basic emotions are not (e.g., Plutchik 1994). If they are not naturally given, they are assumed to result from learning, which often constitutes a social process for human beings. Basic emotions are considered to be innate, whereas non-basic emotions are considered to be learned socially. In this sense, the category of non-basic emotions qualifies as a candidate for social emotions.

2 For an extensive critique of Ekman's experimental work, see Gendron, Crivelli, and Barrett 2018; Russell 1994; 2003.

3 For arguments questioning the dichotomy between basic and non-basic emotions see Scarantino, 2015.

But, first, the claim that emotions can neatly be divided into basic and non-basic emotions has been criticised by many authors (e.g., Clark 2010, Colombetti 2014, Hufendiek 2015). Rebekka Hufendiek, for example, argues that emotions should be individuated as whole-organism responses with reference to their functional profiles (2015). All emotions may have biological aspects that are elaborated through learning (Prinz 2004a). Moreover, the claim that alleged non-basic (and, thus, social) emotions develop later has also been reassessed and criticised (Draghi-Lorenz et al. 2001; Griffiths 2003; Parkinson et al. 2005). Indeed, there may be innate emotions (or emotions that are just as good candidates for innateness as Ekman's six) that have social functions. Some authors have suggested that social emotions can be considered to be basic (e.g., Fessler 2004; Tracy and Matsumoto 2008). Fessler argues that shame—which is typically regarded as a social emotion—is a product of natural selection; i.e., shame is a social emotion, but a rudimentary form of it qualifies as a basic emotion in the sense that it is an evolved human universal (Fessler 2004). If some social emotions are basic and innately rooted, merely defining social emotions via the non-basic criterion, which boils down to the assumption that emotions that are socially learned are non-basic social emotions, is inapt. Indeed, the intuitions driving this objection—e.g., that shame is social—suggest that we have some purchase on the idea of social emotions that does not presuppose an answer to the innateness question. We must therefore consider other candidates for social emotions.

2.2 Higher-Cognitive Emotions

The second candidates for social emotions I'll look at here are higher-cognitive emotions. The term "cognitive" brings to mind theories that claim that cognitive elements—e.g., thoughts or judgments—are constitutive parts of emotions.

The candidate higher-cognitive emotions closely parallel the non-basic emotion proposal, but with greater emphasis on cognition. According to Griffiths (1997) there are some emotions that require more cognitive processing than others. He differentiates between emotions typically considered to be basic emotions—such as anger, fear, and sadness—and higher-cognitive emotions.[4] Griffiths equates the former with Ekman's affect programs, which he sees as innate and automatic. He does not provide a full account of higher-cognitive emotions. Rather, he focuses on pointing out that accounts of emotion that are primarily concerned with basic emotions cannot capture higher-cognitive emotions. Proponents of basic emotions sometimes suggest that non-basic emotions are simple extensions of combinations of basic emotions that require learning (Plutchik 1994). Griffiths thinks this underestimates the differences between basic and other emotions (Griffiths 1997: 102). He argues that higher-cognitive emotions may have long durations, unlike basic emotions, which are short-term, and that many lack stereotypical behavioural display or physiological effects. For example, he puts guilt into this category and says it lacks a facial expression. Most importantly, affect programmes do not do justice to emotions that are integrated with complex cognitive processes (Griffiths 1997).

Griffiths' approach faces several challenges. First, he assumes that some emotions are more cognitive than others, but many other emotion researchers assume that all emotions are equally cognitive. Solomon (1973; 2003), for example, defends an account of emotions as involving evaluative judgments (perhaps unconscious). Nussbaum, similarly, defines emotions as judgments about value and importance (Nussbaum 2001). Likewise for defenders of appraisal theories in psychology (e.g., Arnold 1960; Lazarus 1991; Moors et al. 2013). There are also emotion researchers who assume that all emotions are equally non-cognitive (e.g., Prinz 2004b). Non-cognitivist as well as cognitivist views of emotion aim to encompass all emotions, and not just a specific subset of emotions (such as social emotions).

4 In more recent work, Griffiths has questioned his use of the term "higher-cognitive emotions." For a discussion see Griffiths and Scarantino 2009.

Griffiths has also been criticised for positing an evolutionary divide between basic and higher-cognitive emotions. Focusing on shame, Clark (2010) criticizes Griffiths' distinction by arguing that a lot of emotions that are categorised as higher-cognitive emotions are actually in line with specific basic emotions, and thus could be categorised as basic emotions as well (Clark 2010).

Moreover, if we define social emotions on the basis of the higher-cognitive criterion, we may end up dismissing emotions that carry social elements. So-called basic emotions such as anger might be excluded from the category "social emotions," even in cases where the anger is clearly rooted in social factors (e.g., being angry at a friend for behaving inappropriately in a social scenario).

It may, thus, not be helpful to define social emotions merely via the higher-cognitive emotion criterion.

2.3 Self-Conscious Emotions

As a third candidate, we may consider self-conscious emotions to be social emotions (see Lewis 2016; 2014; Leary 2004; Tangney and Fischer 1995: Tracy and Robins 2004; Wierzbicka 1999). Self-conscious emotions have been said to be central to motivating and regulating people's behaviours, i.e., to contribute significantly to people behaving appropriately according to a given situation (Tangney and Fischer 1995; Tracy and Robins 2004); that is, self-conscious emotions are directly related to norms around socially acceptable behaviours. Social norms determine aptness conditions for emotions. This prescriptive force of social norms in emotions is particularly present in self-conscious emotions; i.e., the connection to social norms via aptness conditions is a necessary criterion for defining self-conscious emotions. According to Leary (2007), "self-conscious emotions are fundamentally social emotions that are elicited by real and imagined events that have potential implications for how the individual is perceived and evaluated by other people" (Leary 2007: 45). In order to ascertain how others perceive and evaluate us, we need self-awareness (cf. Leary 2007: 45). This leads the category "self-conscious emotions" to also be associated with events that seem to threaten our self-evaluation or social status (Hareli and Parkinson 2008; Tracy and Robins 2004). If aptness conditions are harmed, social status is threatened. Here again, self-conscious emotions can be considered directly relevant to social life. This makes them an apt candidate for social emotions.

Let us look at a specific approach to self-conscious emotions, so we can understand how exactly they're linked to sociality. According to Wierzbicka (1999), self-conscious emotions are emotions involving a kind of experience or belief about others' attributions, or meta-cognition about one's actions. For example, you might think that the other person present thinks highly or poorly of you. Or you might self-reflect on a past behaviour or action of yours. Self-conscious emotions can thus be considered a candidate for social emotions, since they involve cognitions about what others think of you or cognitions about your own social behaviour. It is difficult to imagine any of these emotions occurring without being situated in a social scenario (be it imaginary or real). The most common examples of self-conscious emotions are shame, guilt, pride, and embarrassment (Leary 2004; Tracy and Robins 2004; Wierzbicka 1999). Jealousy and admiration are sometimes considered to be examples as well (see, e.g., Bennett and Gillingham 1991; Hareli and Eisikovits 2006; Hareli and Hess 2010).

One problem with this approach is that self-consciousness in relation to emotions seems to entail a broad range of phenomena. Self-consciousness is defined both by referring to the process of thinking about oneself and in referring to the process of thinking about others (see Wierzbicka 1999; Tracy and Robins 2004). For Wierzbicka, "self-conscious" means thinking about other people's judgments of oneself or thinking about one's own actions (Wierzbicka 1999). Tracy and Robins provide an ostensibly similar definition. For them, self-consciousness is constituted by self-awareness and self-representations. "Together, these self-processes

make it possible for self-evaluations, and therefore self-conscious emotions, to occur" (Tracy and Robins, 2004: 105). By entailing such a broad range of phenomena, the demarcation of emotions that are not self-conscious emotions becomes blurry: is self-directed anger not an emotion that results from thinking about one's own actions?

Self-consciousness also fails to address cases of feeling angry or sad or disappointed in social situations where self-consciousness is not at the core of the experience. Consider the example from above again: being angry at a friend for behaving inappropriately in a social situation. This case of anger seems to be social—in the sense that the anger is elicited in and due to a social situation—but not necessarily self-conscious. Likewise when one is sad about being let down or disgusted by hypocrisy. That is, the category 'self-conscious emotions' does not provide a rich account of what 'social' means. This need not be an issue, and indeed might actually help demarcate the border between social and non-social emotions, however. I will come back to this question in Section 3.2.

Self-conscious emotions provide us with a promising candidate for defining social emotions because they are directly connected to social norms around emotions. The emotions considered self-conscious emotions (such as shame, pride, embarrassment, guilt) are also typically categorised as social emotions, as they appear directly relevant to our social everyday lives. But for self-conscious emotions to define the category of social emotions more neatly, the category of self-conscious emotions would have to provide a clearer account of how sociality is connected to self-awareness and self-consciousness.

2.4 Social Function

There are very few explicit efforts to provide a unifying account of social emotions. One is the work of Shlomo Hareli and Brian Parkinson (2008). Hareli and Parkinson (2008) explore the question: What's Social about Social Emotions? While admitting that all emotions are social in some sense, they consider social emotions to be a specific subset of emotions. That is, "social emotions are social in a different way" (Hareli and Parkinson 2008: 131). They argue that, for an emotion to be social, it needs to be necessarily connected to social concerns (Hareli and Parkinson 2008). Social concerns are, for example, concerns about your social status and concerns about what others think of you.

> Shame, embarrassment, and jealousy are social emotions because they necessarily depend on other people's thoughts, feelings or actions, as experienced, recalled, anticipated or imagined at first hand, or instantiated in more generalized consideration of social norms or conventions. Each of these emotions derives its defining quality from an intrinsic relation to social concerns: at the conceptual level, it would not count as a proper instance of the emotion category in question, and at the empirical level it would not have its distinctive relational quality, unless the relevant social concern was in play. (Hareli and Parkinson 2008: 131)

That is, according to this view, social emotions serve social functions. Hareli and Parkinson (2008) argue that tokens of any emotion type may potentially be influenced by social factors or serve social functions, but only "social" emotion types have a necessary and exclusive dependence on social concerns. Hareli and Parkinson's argument is that, in cases where supposedly social emotions operate in non-social ways, non-social events are being appraised as if they were social—e.g., getting angry with one's computer attributes social agency to that non-social entity.

In an earlier approach of emotion and social function, Keltner and Haidt (2001) argue that emotions have social functions and evolve according to these social functions. They differentiate between primordial emotions and elaborated emotions. According to Keltner and Haidt (2001), primordial emotions are composed of physiological, perceptive, and communicative aspects. Elaborated emotions, on the other hand, are "the total package of meanings, behaviors, social practices, and norms that are built up around primordial emotions in actual human societies" (Keltner and Haidt 2001: 11). Keltner and Haidt grant that primordial emotions influence social interactions since they, too, communicate the emotional state of the person expressing the emotion. However, on their view, while primordial emotions influence social interactions, they are not influenced by social interactions in turn.

Claiming that primordial emotions are not influenced by the very social interactions they influence may seem unconvincing, though. For example, the way I feel and express fear may be influenced by the social interactions in which I experience fear. If I was never in a social interaction where somebody had shouted at me before, it might be hard to experience fear when somebody shouts at me for the first time. I might feel estranged instead of afraid. If I am constantly met with praise when I express anger, that may influence the appropriateness conditions I learn about anger.

Once we see this, the "social functions" approach starts to look problematic. It's perfectly true that social emotions serve social functions, but that doesn't help those who want to argue that social emotions are different from other emotions. In response, defenders of these views might say two things: they might say that some emotions are not evolved to be social, while others are; or they might say some emotions are not always social, while others are. I will bypass the evolutionary hypothesis, since evolutionary claims are hard to verify; sociality is part of our evolutionary history, and it would be good to have a criterion about how emotions function now, not just in our evolutionary past.

As for the objection that some emotions are not always social, there is some intuitive pull. For example, one might say that disgust is non-social when applied to rotting food, and social when applied to moral transgressions. Emotions such as embarrassment, in contrast, do seem to have a necessary dependence on social concerns, so we might distinguish these two in this way. Of course, it is not clear that an emotion like embarrassment always has social elicitors; some people might be embarrassed to swim nude even if privacy is guaranteed, and someone can be envious of asteroids that fly across the cosmos. But Hareli and Parkinson account for these facts by arguing that, in cases where social emotions operate in non-social ways, non-social events are being appraised as if they were social (cf. Hareli and Parkinson 2008).

But there is one issue with social function approaches: they seem to assume that there is no way in which food-directed-disgust is social. For Keltner and Haidt (2001), disgust is considered a primordial emotion. And for Hareli and Parkinson (2008) disgust may be considered not to have a necessary relation to social concerns. Now consider the following: the definition of what counts as "rotten food" may depend on culture. This will ultimately also affect when disgust is elicited. Thus, in some sense, disgust may be social. Does this mean disgust is a social emotion? And does this posit a counterargument to distinguishing social from non-social emotions? To analyse this, I will take a closer look at the relationship between sociality and emotions in the next section.

3. Sociality and Emotions

So far, we have seen that there are different candidates for social emotions, each of which defines social emotions via specific criteria. I pointed to some issues with these candidates. Proponents might propose replies, but my review of the candidates is intentionally short in this paper, as it stands to motivate a different question, namely: are all emotions social? To explore this further, I want to pose some questions that might shed light on how to understand emotions to be social in a broader, more generalized sense. Are all emotions impacted by social factors, social norms, and values? And, if all emotions are socially impacted, what does that mean for the nature of emotions? In particular, might it suggest that there is a sense in which all emotions are social?

Before proposing that a pluralistic understanding of social emotions is needed to embrace the different ways in which social emotions may function, this paper needs to engage in a further step: exploring how social factors impact emotions in general. This serves as a basis for explaining what it would mean to propose that all emotions are social. To explore the relation between emotions and sociality in more detail, it is important to turn to the question of how social factors might impact on emotions. Here, I consider it vital to look at social norms, as there is research that points to the involvement of social norms in shaping emotions.

3.1 Social Norms and Emotions

Emotion theorists largely agree that all kinds of emotion are subject to social norms. The issue here is how deeply these norms penetrate into emotions themselves. Ekman (Ekman 1994; Ekman and Cordaro 2011), for example, sees display rules as impacting only on the public manifestation of a pre-existing intact private emotion that itself cannot be seen as social. Social and psychological constructionists take a very different position on the role of social norms and social factors more generally. For example, early constructionists defend an account of emotions as culturally learned patterns of behaviour that have no consistent relationship to biological or physical factors (e.g., Averill 1978; 1980). Before diving deeper into the relationship between social norms and emotions, let me clarify what I mean by social norms here.

We can define social norms as norms about how to comport ourselves in the (real or imagined) presence of others (Hochschild 1979; Averill 1980).[5] An important author to consider when researching social norms is Cristina Bicchieri. According to Bicchieri (2006), social norms can evolve out of conventions or conventional behaviours. Conventions become social norms when they reach sufficient social significance (Bicchieri 2006: 42). This is important to bear in mind when defining social norms as norms about how to comport ourselves in the presence of others: not all rules and guidelines about how to comport ourselves can be considered social norms; some may just be conventions. I will thus enrich this paper's understanding of social norms with Bicchieri's account: social norms are norms about how to comport ourselves in the (real or imagined) presence of others; they are different from social conventions in that they are socially significant. This section cannot provide a full account of what social norms are. Rather, it aims at providing insight into how social norms impact emotions.

5 What I call "social norms" here has been referred to as "social guidelines" or "social rules" by, for example, Hochschild (1979), and "social roles" by, for example, Averill (1980). See e.g. Hochschild (1979: 563): "The social guidelines that direct how we want to try to feel may be describable as a set of socially shared, albeit often latent (not thought about unless probed at), rules." I refer to these factors as social norms, since I consider this to be a term that encompasses guidelines, rules, and roles.

In the following, I will look at a kind of social norm that has been explicitly addressed in emotion research: gender norms. The socially built categories of binary gender have enormous consequences on our construction as individuals, and on the broader organisation of society (Butler 1990; Ahmed 2006). Gender stereotypes are transmitted very efficiently through socialisation and education. This affects our emotional development and thus also influences how we display and understand emotions. Let's look at some examples to see how gender norms influence our emotional life.

Research on blushing, for example, has shown that more women than men tend to blush (Eickers 2022; Crozier 2006; Darby and Harris 2013). This does not point to women being more prone to blushing by nature. Rather, women are more likely to be taught to be shy, apologetic, and adaptive than men—who are, on the other hand, more likely to be taught to be demanding, entitled, and even aggressive.[6] It has been shown that men are more likely to express emotions associated with power in general (e.g., anger, disgust, pride), whereas women are more likely to express emotions associated with weakness (e.g., sadness, fear, shame) (Fischer and Manstead 2000; Fischer et al. 2004). That is, gender norms have an impact on our emotional lives, and specifically on how we experience and express emotions (see Brody and Hall 2010 for an extensive analysis). Gender norms are not the only social norms that influence our emotions in this way: there are also social norms and expectations around race and emotions (Leboeuf 2017), as well as social status and emotions (Alexander and Wood 2000; Parkinson et al. 2005), for example.[7]

Looking at how social norms play a role in how emotions evolve, how emotions are expressed and communicated, and how emotions are recognised, can show us that the social does not only affect emotions that are categorised as social but will also affect basic emotions and emotions that are considered by some to be innate or universal, e.g. anger or fear, as shown above, as well as joy (think of cultures that denounce hedonism) and sadness (consider norms against men crying). It is hard to think of any emotion untouched by such norms. Likewise, the way we express emotions can be impacted by norms pertaining to age, ethnicity, class, and so on.

Now, if social norms impact even so-called basic emotions, might that mean that the category "social emotions" becomes obsolete because, on this understanding, all emotions are social? In the following, I will answer that question.

3.2 Are All Emotions Social?

Section 2 surveyed candidates for social emotions, meaning: different accounts and criteria that have been referred to in emotion theory as, or which can be considered to be, social emotions. Through brief criticisms of the candidates for social emotions, Section 2 motivated us to embrace a pluralistic understanding of social emotions. Section 3 so far has provided us with reasons to think that social norms impact emotions in general—which now poses the question of how to distinguish social emotions from non-social emotions. In the following, I will be concerned with asking whether, as a consequence, we need to think that all emotions are social.

6 For a detailed discussion of gender and emotion see Fischer 2000 and Shields 2000.

7 "Research by Tiedens and colleagues (Tiedens, 2001; Tiedens, Ellsworth, and Mesquita, 2000) has shown that high-status individuals are expected to respond with anger (rather than sadness or guilt) to negative outcomes and with pride (rather than appreciation) to positive outcomes" (Van Kleef 2016: 74).

Social and psychological constructionist accounts of emotion may lead our way to finding an answer to this question. Constructionists such as Hochschild (1979), Averill (1980), Lutz and White (1986), Mesquita et al. (2016), Parkinson et al. (2005), and Barrett (2009; 2017) provide accounts of emotion that are closely connected to social norms. Constructionists propose that human emotions are not simply innate responses, even if biology makes a contribution; rather, they argue that the way we emote is impacted by socialisation. For example, social factors can influence what elicits our emotions, how we conceptualise them, how we express them, how negative or positive they seem, and the actions they potentiate. Bear in mind that constructionist approaches to emotion also differ: Barrett defends an account of how concepts and language set up our emotion categories that refers to Searle's concepts of collective intentionality and collectively ascribing functions: "A physical event like a change in heart rate, blood pressure, or respiration becomes an emotional experience only when we, with emotion concepts that we learned from culture, imbue the sensations with additional functions by social agreement" (Barrett 2017: 39). Other accounts, such as those of Parkinson et al. (2005) or Hufendiek (2020), do not focus on concepts and language, but rather on the role of social structure and social practices in constituting emotion types. Analogously, Mesquita and Parkinson (in press) distinguish a social concept and a social role approach that also helps to demarcate theoretical differences among constructionists.

What social constructionists have in common, though, is that they argue that it is the way that emotions are affected by social norms and social factors that gives us reason to believe that emotions are socially constructed. This may seem like a radical view, but it shouldn't be regarded as such. As the example of gender norms show, social influence is indeed pervasive. In fact, authors such as Ekman (1999; Ekman and Cordaro 2011) grant as much, saying that culture can influence how emotions are expressed. For Ekman, expression is part of the emotion, so his concession allows a degree of social construction. There are more radical views, of course, which downplay biological contributions, but the point I am making here applies even if we say that all human emotions are constructed (that is to say, socially influenced) to some degree. In a sense, this might mean that all emotions are social, since emotions in general are socially impacted, not just social emotions. Let us look at examples of how constructionists argue for the claim that all emotions are, in some sense, social.

Lutz and White (1986) point out that emotions are linked to cultural and individual factors since these are all part of the social structure in one way or another. They identify different ways in which emotions are related to social structure. These are: "The degree of individualism, notions of privacy and autonomy, multiplicity of selves, or sense of moral responsibility" (Lutz and White 1986: 420). That is, Lutz and White also draw on connections between the social and the emotional, and essentially state that all emotions are constructed by different aspects of the social structure.

Similarly, Parkinson et al. (2005) analyse emotions with respect to their social relations. They look at how social groups affect emotions, for example. Expressions and feelings of anger, sadness, and fear can be mitigated by in-group and out-group effects (Parkinson et al., 2005). For example, we are more prone to feel and show fear towards someone we perceive as not being a member of our own social group than we are towards a member of our own social group. Leaving in- and out-group effects aside, there are also studies that indicate we are more likely to smile when watching a film with friends than when we are alone (Parkinson et al. 2005). When discussing social emotions specifically, Parkinson et al. emphasise that "even so-called 'basic' emotions have interpersonal origins and develop in close attunement with social relations" (Parkinson et al. 2005: 188), pointing out that all emotions are social in the sense that they are part of social interactions and relations.

Defences of constructionist perspectives on emotions can also be found in more recent work. Mesquita et al. (2016), for example, argue that emotions are constructed through cultural norms and that culturally

normative emotions, in turn, serve to maintain social relationships (within that culture). “The combined research on cultural differences suggests that emotions emerge through processes of construction. . . . Emotions are iterative and active constructions that help an individual achieve the central goals and tasks in a given (cultural) context” (Mesquita et al. 2016: 34). Is any emotion immune to social influences such as cultural norms? Consider disgust at rotten food again. Cultures differ in attitudes towards decay and edibility. Icelanders, for example, eat putrefied fish, which people from other countries may find disgusting. What we eat, how we eat, and how we express aversion are all socially impacted.

All in all, constructionists give us multiple reasons to think that social factors do not only have an impact on emotions we deem to be social in the first place but on all kinds of emotions. Even those who don’t like to call themselves constructionists must heed that lesson from constructionist theories. If we grant that some emotions are innate, for example, it still seems clear that they can be impacted by social norms. This may be interpreted as entailing that all human emotions are social in some sense.[8]

Does embracing that all emotions are social in some sense mean that all emotions are social emotions? The short answer is no. First, in many contexts, it is helpful to distinguish social and non-social emotions: food disgust and moral disgust may differ in that the latter is ultimately directed toward something social. Second, this points to the importance of differentiating a specific social emotions category from other emotion categories, and of understanding social emotions pluralistically. There may be a difference between “basic disgust” and “social disgust”; but we may also see, here, how social functions and self-awareness are important for building a social emotions category.

Section 2 has equipped us with reasons to be sceptical about trying to come up with a clear definition of social emotions and trying to distinguish social emotions from non-social emotions via one specific criterion. All of the candidates for social emotions we have encountered face the issue of excluding or ignoring cases of emotion that may well be deemed social but are not accounted for by the conditions of the particular candidate. Furthermore, the candidates do not offer a clear explanation of the connection between sociality and emotions. This can be accounted for by a pluralistic account that considers multiple social factors to be defining criteria for a social emotions category.

That is, the question “Are all emotions social?” may be answered in the following way: we have good reason to believe that all emotions are social, in the sense that all emotions are socially impacted, but we do not have sufficient reason to believe that all emotions belong to a distinct social emotions category.

4. Conclusion

This paper has looked into different candidates for social emotions and briefly assessed them. This has motivated embracing a pluralistic understanding of social emotions. It has also motivated going further, prompting us to ask whether all emotions are social.

8 The question I leave aside here is whether, and in what sense, non-human emotions are (always) social. For some animals, the answer may be that they are not, but it is noteworthy that many animals have rich social lives (including interactions with us), and all are capable of learning throughout their lifespans.

While Section 2 focused on analysing different candidates for social emotions, Section 3 proceeded to move the focus from looking at different candidates for the category social emotions to asking 'Does the social impact emotions in general?' I specifically looked at social norms and how they affect emotions, and then asked whether all emotions are social. While each of the candidates reviewed in Section 2 implies that not all emotions are social, Section 3 provided us with reason to think that all emotions are social in some sense. But Section 3 also asked whether this means that all emotions are social emotions and answered that question, saying that not all emotions are social emotions. Saying that all emotions are social in some sense doesn't jeopardize the category social emotions as such. How exactly, and to what degree, different emotions are socially impacted is an open question for research. Emotions may be social in a variety of different ways.

Exploring the connection between sociality and emotion further points towards embracing a pluralistic account of social emotions. The implication of this final insight is not that we should abandon the possibility of any social vs. non-social emotion distinction, but rather to seek to be more precise in delineating different varieties and instances of social emotions.

References

Ahmed, S. 2006. Queer Phenomenology: Orientations, Objects, Others. Duke University Press.

Alexander, M. G. and W. Wood. 2000. "Women, Men, and Positive Emotions: A Social Role Interpretation." In Gender and Emotion: Social Psychological Perspectives, edited by A. H. Fischer, 189–210. Cambridge University Press.

Arnold, M. B. 1960. Emotion and Personality. New York: Columbia University Press.

Averill, J. R. 1978. "Anger." In Nebraska Symposium on Motivation, edited by H. Howe and R. Dienstbier, 1-80. University of Nebraska Press.

———. 1980. "A Constructivist View of Emotion." In Emotion: Theory, Research and Experience. Vol. 1: Theories of Emotion, edited by R. Plutchik and H. Kellerman, 305–40. Academic Press.

Barrett, L. F. 2009. "Variety is the Spice of Life: A Psychological Construction Approach to Understanding Variability in Emotion." Cognition & Emotion 23: 1284–306.

———. 2017. How Emotions are Made. Houghton Mifflin Harcourt.

Barrett, K. C. and J. J. Campos. 1987. "Perspectives on Emotional Development II: A Functionalist Approach to Emotion." In Handbook of Infant Development, 2nd ed., edited by J. Osofsky, 555–78. Wiley.

Bennett, M., and K. Gillingham. 1991. "The role of self-focused attention in children's attributions of social emotions to the self." The Journal of Genetic Psychology, 152, 303–309.

Bicchieri, C. 2006. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press.

Brody, L. R. and J. A. Hall. 2010. "Gender, Emotion, and Socialization." In Handbook of Gender Research in Psychology. Volume 1: Gender Research in General and Experimental Psychology, edited by J. C. Chrisler and D. R. McCreary, 429–51. Springer.

Buck, R. 1999. "The Biological Affects: A Typology." Psychological Review 106: 301–36.

Butler, J. 1990. Gender Trouble: Feminism and the Subversion of Identity. Routledge.

Clark, J. A. 2010. "Relations of Homology between Higher Cognitive Emotions and Basic Emotions." Biology & Philosophy, 25(1): 75–94.

Colombetti G. 2014. The Feeling Body: Affective Science Meets the Enactive Mind. Cambridge, MA: MIT Press.

Crozier, W.R. 2006. Blushing and the Social Emotions. The Self Unmasked. Palgrave Macmillan.

Darby, R. S. and C. R. Harris. 2013. "A Biosocial Perspective on Embarrassment." In The Psychological Significance of the Blush, edited by W. R. Crozier and P. J. de Jong, 120–46. Cambridge University Press.

Darwin, C. 1872. The Expressions of Emotions in Man & Animals, 1st ed. Philosophical Library.

Draghi-Lorenz, R., V. Reddy, and A. Costall. 2001. "Re-Thinking the Development of 'Non-Basic' Emotions: A Critical Review of Existing Theories." Developmental Review 21: 263–304.

Ekman, P. 1972. Emotions in the Human Face. Pergamon Press.

———. 1994. "All Emotions are Basic." In The Nature of Emotion: Fundamental Questions, edited by P. Ekman and R. J. Davidson, 15–19. Oxford University Press.

———. 1999. "Basic Emotions." In The Handbook of Cognition and Emotion, edited by T. Dalgleish and T. Power, 45–60. John Wiley & Sons.

Ekman P. and D. Cordaro. 2011. "What is Meant by Calling Emotions Basic." Emotion Review 3: 364–70.

Ekman, P. and W. V. Friesen. 1971. "Constants across Cultures in the Face and Emotion." Journal of Personality and Social Psychology 17: 124–9.

Fessler, D.M.T. 2004. "Shame in Two Cultures: Implications for Evolutionary Approaches." Journal of Cognition and Culture 4(2): 207-262.

Fischer, A. H., ed. 2000. Gender and Emotion: Social Psychological Perspectives. Cambridge University Press.

Fischer, A. H. and A. S. R. Manstead. 2000. "The Relation between Gender and Emotion in Different Cultures." In Gender and Emotion: Social Psychological Perspectives, edited by A. H. Fischer, 71–94. Cambridge University Press.

Fischer, A. H., P. M. Rodriguez Mosquera, A. E. M. van Vianen, and A. S. R. Manstead. 2004. "Sex and Culture Differences in Emotion." Emotion 4: 87–94.

Gendron, M, C. Crivelli, and L. F. Barrett. 2018. "Universality Reconsidered: Diversity in Making Meaning of Facial Expression." Current Directions in Psychological Science 27 (4): 211–9.

Griffiths, P. E. 1997. What Emotions Really Are: The Problem of Psychological Categories. University of Chicago Press.
———. 2003. "Emotions." In The Blackwell Guide to the Philosophy of Mind, edited by S. Stich and T. Warfield, 288–308. Blackwell.
Griffiths, P. E. and A. Scarantino. 2009. "Emotions in the Wild: The Situated Perspective on Emotion." In Handbook of Situated Cognition, edited by P. Robbins and M. Ayede, 437–54. Cambridge University Press.
Hareli, S. and B. Parkinson. 2008. "What's Social about Social Emotions?" Journal for the Theory of Social Behaviour 38: 131–56.
Hareli, S. and Z. Eisikovits. 2006. "The Role of Communicating Social Emotions Accompanying Apologies in Forgiveness." Motivation and Emotion 30: 189–97.
Hareli, S. and U. Hess. 2010. "What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception." Cognition and Emotion 24: 128-140.
Hochschild, A. R. 1979. "Emotion Work, Feeling Rules, and Social Structure." American Journal of Sociology 85: 551–73.
Hufendiek, R. 2015. Embodied Emotions. A Naturalist Approach to a Normative Phenomenon. Routledge Studies in Contemporary Philosophy.
———. 2020. "Emotions, Habits, and Skills." In Habits. Pragmatist Approaches from Cognitive Science, Neuroscience, and Social Theory, edited by F. Caruana and I. Testa, 100–19. Cambridge University Press.
Izard, C. E. 1977. Human Emotions. Plenum Press.
———. 2011. "Forms and Functions of Emotions: Matters of Emotion–Cognition Interactions." Emotion Review 3: 371–8.
Keltner, D. and J. Haidt. 2001. "Social Functions of Emotions." In Emotions: Current Issues and Future Directions, edited by T. Mayne and G. A. Bonanno, 192–213. Guilford Press.
Lazarus, R. S. 1991. Emotion and Adaptation. Oxford University Press.
Leary, M. R. 2004. "Digging Deeper: The Fundamental Nature of 'Self-Conscious' Emotions." Psychological-Inquiry 15: 129–31.
———. 2007. "How the Self Became Involved in Affective Experience: Three Sources of Self-Reflective Emotions." In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), The self-conscious emotions: Theory and research (pp. 38–52). The Guilford Press.
Leboeuf, C. 2017. "Anger as a Political Emotion: A Phenomenological Perspective." In The Moral Psychology of Anger, edited by M. Cherry and O. Flanagan, 15–30. Rowman and Littlefield.
Lewis, M. 2014. The Rise of Consciousness and the Development of Emotional Life. Guilford Press.
———. 2016. "Self-Conscious Emotions: Embarrassment, Pride, Shame, Guilt, and Hubris." In Handbook of Emotions, 4th ed., edited by L. F. Barrett, M. Lewis, and J. M. Haviland-Jones, 792–814. Guillford Press.
Leys, R. 2017. The Ascent of Affect: Genealogy and Critique. University of Chicago Press.
Lutz, C. A. and G. White. 1986. Annual Review of Anthropology: The Anthropology of Emotion. Annual Reviews.
Mesquita, B., M. Boiger, and J. De Leersnyder. 2016. "The Cultural Construction of Emotions." Current Opinion in Psychology 8: 31–6.
Mesquita, B. and B. Parkinson. In press. "Social Constructionist Theories of Emotions." In Emotion Theory: The Routledge Comprehensive Guide, edited by A. Scarantino. Routledge.
Moors, A., P. C. Ellsworth, K. R. Scherer, and N. H. Frijda. 2013. "Appraisal Theories of Emotion: State of the Art and Future Development." Emotion Review 5 (2): 119–24.
Nussbaum, M. C. 2001. Upheavals of Thought: The Intelligence of Emotions. Cambridge University Press.
Panksepp, J. 2007. "Neurolizing the Psychology of Affects: How Appraisal-Based Constructivism and Basic Emotion Theory Can Coexist." Perspectives on Psychological Science 2: 281–96.
Parkinson, B., A. H. Fischer, and A. S. R. Manstead 2005. Emotion in Social Relations: Cultural, Group, and Interpersonal Processes. Psychology Press.
Plutchik, R. 1994. The Psychology and Biology of Emotion. Harper Collins College.
Prinz, J. J. 2004a. "Which Emotions Are Basic?" In Emotion, Evolution, and Rationality, edited by D. Evans and P. Cruse, 69–88. Oxford University Press.
———. 2004b. Gut Reactions. A Perceptual Theory of Emotion. Oxford University Press.

Russell, J. A. 1994. “Is There Universal Recognition of Emotion from Facial Expression: A Review of the Cross-Cultural Studies.” Psychological Bulletin 115: 102–41.

———. 2003. “Core Affect and the Psychological Construction of Emotion.” Psychological Review 110: 145–72.

Scarantino, A. 2015. “Basic Emotions, Psychological Construction, and the Problem of Variability.” In The Psychological Construction of Emotion, edited by L. F. Barrett and J. A. Russell, 334–76. Guilford Press.

Shields, S. A. 2000. “Thinking about Gender, Thinking about Theory: Gender and Emotional Experience.” In Gender and Emotion: Social Psychological Perspectives, edited by A. H. Fischer, 3-23. Cambridge University Press.

Solomon, R. 1973. “Emotions and Choice.” The Review of Metaphysics 27: 20–41.

———. 2003. Thinking About Feelings: Philosophers on Emotions. Oxford University Press.

Tangney, J. P. and K. W. Fischer. 1995. Self-Conscious Emotions: The Psychology of Shame, Guilt, Embarrassment, and Pride. Guilford Press.

Tiedens, L. Z. 2001. “Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral.” Journal of Personality and Social Psychology 80: 86–94.

Tiedens, L. Z. and P. C., Ellsworth and B. Mesquita. 2000. “Sentimental stereotypes: Emotional expectations for high- and low-status group members.” Personality and Social Psychology Bulletin 26: 560–575.

Tracy, J. L. and D. Matsumoto. 2008. “The spontaneous expression of pride and shame: evidence for biologically innate nonverbal displays.” Proceedings of the National Academy of Sciences of the United States of America 105(33): 11655–11660.

Tracy, J. L. and R. W. Robins. 2004. “Putting the Self into Self-Conscious Emotions: A Theoretical Model.” Psychological Inquiry 15: 103–25.

van Kleef, G. 2016. The Interpersonal Dynamics of Emotion: Toward an Integrative Theory of Emotions as Social Information. Cambridge University Press.

Wierzbicka, A. 1999. Emotions across Languages and Cultures: Diversity and Universals. Cambridge University Press.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Received: 15 November 2023
Accepted: 27 February 2024
Published: 30 June 2024

To crave is like to be in love[1]

Rachel Frenette – Paris 1 Panthéon-Sorbonne (philosophy department), France

Abstract

In this article, I show that the experience of addictive desires, which I also refer to as cravings, is similar to the experience of the state of being in love or of *limerence*. In other words, I argue that some part of the experience of addiction resembles some part of the experience of romantic love. Many in the literature have tried to show that one can be addicted to love in the way that love constitutes an addiction. Yet, none seem to have taken the opposite route, which takes love not as an object of addiction, nor as something that can be reduced to it, but rather as a model from which we can understand addiction. This is odd, as such a perspective is warranted by first-person accounts of addiction. Therefore, in my view, *addiction is like love*, and not the other way around. Nevertheless, it is not addiction as a whole that is similar to love; love in itself cannot be compared to addiction: it is precisely the experience of addictive cravings in addiction that is like limerence. In order to put forward the similarities between experiences of cravings and experiences of limerence, I go through the basic components of limerence as described by Dorothy Tennov (1979), and prove that they are found in the experience of cravings. This comparison between the two phenomena allows me to introduce the notion of reciprocation in the addiction literature.

1. Introduction

Interestingly, many recovered addicted individuals refer to their past drug use through the lexicon of love, romance, and relationships. Take for example Caroline Knapp (1999), an American columnist, who titled her now well-known memoir on alcoholism *Drinking: A Love Story*. Similarly, the neuroscientist Judith Grisel talks about her addiction to pot like a loving friendship: “From the first time I got high until long after I’d smoked my last bowl, I loved marijuana like a best friend. This is not hyperbole” (2020, 51). This way of speaking is somewhat puzzling, because love and addiction have, intuitively, opposite valences: on one hand, love is thought to be an overall positive thing; on the other hand, the condition of addiction holds negative value. Yet if recovered addicted individuals intentionally use these terms and explicitly choose to talk about their experience of addiction through the vocabulary of love, we may wonder what it is exactly about addiction that makes it appear similar to love. For example, would it be accurate to say that addicted individuals are *in love* with a substance? Is

1 **Acknowledgements**: Research for this article was supported by the FRQSC in Canada (research grant 321905). I am also thankful to Matt Anderson as well as both anonymous reviewers for their helpful and thoughtful comments.

this way of speaking purely metaphorical, or does it rather express a certain truth about the relation held between an addicted person and the object of their addiction?

In this article, I wish to show that the experience of addictive desires is similar to the experience of limerence, of the state of being in love.[2] Many in the literature have tried to show that one can be addicted to love in the way that love constitutes an addiction (Peele and Brodsky 1975; Timmreck 1990; Reynaud et al. 2010; Sussman 2010; Lewis 2015; Fischer et al. 2016, Earp et al. 2017). Yet, none seem to have taken the opposite route, which takes love not as an object of addiction, nor as something that can be reduced to addiction, but rather as a model which we can use, by way of comparison, to understand addiction. In this regard, *addiction is like love*, and not the other way around. Nevertheless, it is not addiction as a whole that is similar to love; love in itself cannot be compared to addiction. Rather, the experience of addictive cravings in addiction is like *limerence*, which is the technical term given to the state of being in love.

In order to put forward these similarities, I will go through the basic components of limerence as described by Dorothy Tennov, the psychologist who coined the term in *Love and Limerence* (1979), and prove that they are found in the experience of cravings. I hope to show with this article that there is something highly interesting to note in the features that cut across both the state of being in love and the state of being addicted through the experience of addictive cravings. Love is a word that has an equivalent in every language, unlike craving or addiction (Hormes and Rozin 2010). It might therefore be more universal and cross-cultural than addiction. If that is the case, then it seems better suited to try to make sense of addiction through what we know of love and relationships, rather than the other way around.

2. What is Limerence?

“Limerence” is a term coined by Dorothy Tennov, as she was trying to differentiate the act of loving and the state of being romantically in love. According to Tennov, the first is something you do, and consequently choose, while the second is a state, a mental activity (1979, 18), and therefore is rather endured and felt. The distinction between the two implies that you can love without being in love, and be in love without loving. Notably, both are components of romantic love. However, the first is a non-necessary component of it, because the phenomenon of romantic love does not require that one *be in love*; one can love romantically without being in love as characterised by limerence[3]. Sexual desire must also be distinguished from love and limerence. Although they are strongly correlated, the three can exist without each other. Another state that limerence must be differentiated from is infatuation. Although Robert Nozick’s (1989) description of infatuation[4] seems to fit quite well with Tennov’s account of limerence, Lopez-Cantero (2022) argues that infatuation, unlike limerence, is an *unreasonable* state. Because the experience of cravings is one that seems to be typically *arational*,[5] in the sense that it is neither reasonable nor unreasonable, I think it is better suited for comparison with limerence, and not infatuation.

2 I want to stress that the state of being in love and the act of loving are different phenomena. What I compare with the experience of cravings is precisely the experience of *being in love*. One of the differences between being in love and loving is that the first is a state, not an act, and can therefore be (uncontroversially) self-centred and self-interested. This is so because being in love does not require that anyone except oneself be involved in such a state, outside imagination. The fact that being in love and loving are different phenomena seems to be corroborated by first-person testimonies, like in Tennov’s work, but also by literature and cinema. Take for example the beginning of the movie *Saltburn*, where the character Oliver narrates the following: “I wasn’t in love with him. I know everyone thought I was. But I wasn’t. I loved him. Of course!” (Fennell 2023).

3 This is something that Tennov herself recognises.

4 “Infatuation” is a synonym of “being in love” in his book (see Nozick 1989, 70).

5 This is so because cravings do not need to be triggered by (good or bad) reasons. Most often they can be rationalised *a posteriori*, but no reason is needed for them to arise.

Tennov is not alone in thinking that the state of being in love corresponds to what she identifies as limerence. In fact, a study conducted by Lamm and Wiesmann (1997) shows how people's experiences of the state of being in love include many of the characteristics listed by Tennov as basic components of limerence, and which I will explore in section 4, such as "arousal," "attributing positive characteristics to O," and "thinking about O," where O is the object of the love. Moreover, José Ortega y Gasset's description of love's initial stage (before the act of love) seems to correspond closely to what Tennov describes as being in love (1957, 12):

> At its inception love certainly resembles desire, because its object, whether person or thing, excites it. The soul feels agitated, delicately wounded in one spot by a stimulus produced by the object. Such a stimulus has, then, a centripetal direction: it comes to us from the object. But the act of love does not begin until after that excitement, or rather, incitement.

In other words, there is arousal and interest towards the loved object, for Ortega y Gasset, at the inception of love; a pulling effect on attention and perception from the loved object. Such a definition of the phenomenon of love at its start is very similar to limerence. At least, and as we will see in section 4 of this article, both share important common traits.

3. What Is a Craving?

As Weingarten and Elston (1990) describe it, most people have an intuition about the meaning of the word *craving*. Usually, this intuition amounts to the idea that a craving is a subjectively intense desire for something (Kozlowski and Wilkinson 1987; Bruehl et al. 2006).[6] But of course, while many agree on this view,[7] different theories have different explanations of what cravings are and of how they are formed. For example, Dill and Holton (2014) and Henden (2008) build from the theory of incentive sensitisation (Robinson and Berridge 1993) in order to explain the cause of cravings, which they understand to be intense desires to consume certain drugs. For Loewenstein (1999), cravings are not rooted in incentive sensitisation processes, but are rather comparable to visceral factors because of their effect on perception and attention. In this way, cravings are similar to hunger, thirst, or pain, which are motivational and appetitive drives normally essential to our survival. Of course, in the case of addiction, cravings for drugs are not essential for our survival, and for Loewenstein this makes them irrational. Others, like Foddy and Savulescu (2007), take cravings to be part of the larger category of pleasure-oriented desires. Therefore, for these authors, cravings are not pathological and are rather at the extreme end of the continuum of pleasure-oriented desires.

Overall, theories of cravings center around a few questions we may ask ourselves about their characterisation, whether they be food or drug cravings, which are the following: 1) What is a craving? 2) What triggers a craving? 3) What is the relationship between craving and consumption? Type 1 questions want to understand what cravings are, ontologically speaking. Questions about the distinction between cravings and ordinary desires go in this category. Type 2 questions study the cause of the craving in the moment when it is at its full effect. For example, the cause may be internal and rooted in specific brain mechanisms, or it may be external, through the presence of some environmental cues. Type 3 questions are concerned with the relationship

6 The object of the craving is not always the same for different authors. For example, Marlatt (1978) claims that, in the case of alcohol, it is the effects of alcohol rather than the alcohol itself that are desired.

7 If intensity of cravings is usually agreed upon, Sripada (2018) argues that cravings are not really intense, but rather very *frequent* desires to consume, which makes them incredibly hard to resist by the principle of fallibility.

between a craving and the effective act of consumption: is it causal, or simply correlational? Some believe that there is a direct link between craving and consumption and relapses, while others hold that there is no such relation between the two (Tracy 1994).

But there are also Type 4 questions, which are not as present in the literature, and yet are equally as important. These questions focus on the experience one has of a craving. The experience of a craving is not independent of its cause or effects. It is also indirectly related to ontology. Nevertheless, it allows for a different approach in understanding cravings, focused on the moment when they are felt and on the subjective perception one has of them. In this article, I wish to address Type 4 questions and matters of perception and phenomenology. In other words, in order to make a parallel between cravings and limerence, I want to know how cravings are experienced.

In my view, a craving can be characterised through four main features: it is intentional; it is accompanied by a feeling of urgency; it involves mental imagery; it is goal-directed.

The first feature is rather evident: a craving is a craving for *something* (Elster 1999). Usually, this object is specific and determined for the person (alcohol for an alcoholic, cigarettes for a smoker), but the object can be loosely determined, meaning that a *type* of object can be the object of the craving. For example, a person can crave opiates in general, as a type, and therefore want to consume a variety of objects that fall under the type "opiates" (Dhawan et al. 2002).

The second feature[8] highlights the fact that the experience of a craving is one that is intense, almost overwhelming. Urgency happens in two ways: first, a craving person feels that he must attend to his craving *now*; second, there is urgency in the fear of never feeling satisfied and satiated. In other words, urgency is active, in the way that it is action-forward, and it is affective because it also involves a fear of being without the substance, of not having the possibility before one to consume.

The third feature reveals that most experiences of cravings involve mental imagery, so that people experiencing cravings for a certain substance will imagine themselves with the substance or imagine the substance itself in very vivid and visual ways (Salkovskis and Reynolds 1994; Tiffany and Drobes 1990; Andrade et al. 2012).

And finally, the fourth feature addresses the idea that having a craving very often leads to an act. In the case of addiction, this goal-directedness is specifically understood as leading to the act of using drugs. This last feature is explicable by the Elaboration-Intrusion theory of desire (Andrade et al. 2012). If a craving is an intrusive thought which has been elaborated, meaning that plans have been made in order to satisfy the desire, then it seems as if a craving is similar to a goal one wants to achieve.

8 Many quotes from addiction memoirs point towards this idea of the two-fold urgency. Take for example this passage in Caroline Knapp's memoir: "A woman I know named Liz calls alcoholism 'the disease of more,' a reference to the greediness so many of us tend to feel around liquor, the grabbiness, the sense of impending deprivation and the certainty that we'll never have enough. More is always better to an alcoholic; more is necessary" (1999, 57).

4. The Basic Components and the Link with Cravings

In her book, Tennov (1979, 23–24) lists twelve basic components[9] that characterise the state of limerence. In this section, I will go through each of the components, and show that they can be found, in a similar fashion, in descriptions of experiences of cravings in addiction.

a. Intrusive thinking about the object of your passionate desire (the limerent object or "LO"), who is a possible sexual partner.

Tennov stresses that limerence is "first and foremost a condition of cognitive obsession" (1979, 33). In other words, limerence manifests itself predominantly in the way that it interferes with thoughts. The limerent object ("LO") is present in the mind intrusively and persistently when limerence is at its peak intensity. The limerent fantasy is not something that you can control, as it merely happens to you (1979, 40). In this way, it is considered almost inescapable. Tennov cites a person's diary in order to show how the intrusiveness of limerence is experienced: "This obsession has infected my brain. I cannot shake those constantly intruding thoughts of you. Every thought winds back to you no matter how hard I try to direct its course in other directions" (1979, 34). For a limerent person, therefore, everything is intrusively and obsessively reminiscent of the LO, whether it be a thing, a place, a person, or an event.

It appears that this notion of intrusive thinking is also very central in the case of the experience of cravings. In order to understand the link between thought-intrusiveness and cravings, we can refer ourselves to the Elaborated-Intrusion theory first developed by Kavanagh et al. (2005). This theory of desires, which takes cravings to be a specific type of desires in general, holds that a desire is an intrusive thought that eventually turns into a process of cognitive elaboration. An initial intrusive thought can be triggered by cues of various sorts and of internal or external origins. Once an intrusive thought possesses enough affective weight, it can be elaborated, and it is only then that it acquires the status of a desire. The elaboration consists in "planning ways of achieving the desire ('I could buy a coffee'), generating expectancies about satisfying the desire ('I'd be able to concentrate better once I've had a coffee'), and thinking about one's self-efficacy or ability to obtain the desired object or activity ('I've got change for the coffee machine in the lobby')" (Andrade et al. 2012, 130). To put it briefly, a craving is, according to this theory, an *intrusive thought* which holds an affective weight and which becomes a desire through cognitive elaboration.

It should be pointed out here that in limerence Tennov claims that there is something about the *sexual appeal* of LO that fosters being in love and romantic attraction. In fact, in limerence, for an LO to be an intrusive LO, it must be considered as sexually appealing by the person in the limerent state.[10] Of course, in substance addictions, the object of an addiction cannot have such a status. How, then, can we make sense of this part of the first element in the case of the experience of cravings?

There are two possible ways to solve this tension: either we redefine the element of sexual appeal in a way that can be integrated and applied in the case of addiction, or we accept that there is no such element in the case of the experience of cravings and explain how it is not a problem for a comparison. I suggest taking the first path, and therefore redefining the element in a way that maintains what Tennov is trying to explain in the case

9 Tennov does not mention that the elements are necessary and sufficient components of limerence. Therefore, I think it fair to understand them as *prototypical* elements of the experience of limerence. As a consequence, although her account might appear as very normative, it does not actually state strict necessary and sufficient aspects of limerence.

10 Tennov recognises how this element can be absent in limerence, therefore acknowledging the possibility that asexual people experience limerence nevertheless.

of limerence, but which can also be suited for making sense of the experience of cravings. In this way, when Tennov says that for an LO to be an LO he or she needs to be sexually appealing, she really means that for an LO to be an LO he or she needs to be *seen as an object that can be rewarding*. Sex is a natural reward, and so it could be the case that what really influences any romantic pull towards a person in particular is the perception that they can be rewarding, or bring about a reward, in the future. In the specific case of limerence, the reward is taken as something that is attained through sex or through a possible sexual encounter. If this reductionist strategy is correct, then it is arguable that this element is present in the case of addictive cravings too. In fact, for an object to be a craved object (a "CO"), it needs to be seen as an object that can be rewarding or that can eventually bring about a reward. In addiction, the reward will be obtained differently than through a sexual encounter. Nevertheless, the CO, just like the LO, will be perceived as rewarding.

b. Acute longing for reciprocation.

When one is in a state of limerence, one deeply wishes for reciprocation from his LO. One wants to feel loved back and seeks for proof of reciprocation everywhere. The hope that reciprocation is possible and eventually going to happen keeps the person in the romantic state. Tennov cites the experience of a woman named Hilda, which accounts very well for this second element (1979, 60):

> Actually, all the time Stu and I were lovers, it was like that. Even when we were together every week without fail, I'd be consumed with hopes and plans and visions of him really showing love, and he'd always pull something unexpected. If I gave up and planned to break it off, that would be the weekend he'd start off with flowers 'just for my Hilda' and be especially nice. The next week when I'd work myself up to expecting a big response from him, there would be the inevitable letdown.

Hilda's experience of being in love with Stu really highlights the fact that when one is in love, one is consumed with the thought of the other person reciprocating love, presence, and affection. Hilda in this example is, even through adversity and plans gone wrong, lost in her thoughts and fantasies that Stu might, eventually, reciprocate his love for her.

Reciprocation in the case of addiction is, of course, not to be assumed to be reciprocation of *love*. Also, if reciprocation is defined as something necessarily done by an agent, then it cannot be accounted for in substance addiction. Yet, despite these differences, I believe that there is something in the experience of cravings that resembles reciprocation if we define the latter, going back to its Latin etymology, as a movement of back and forth, or more specifically as a kind of giving back, a *return of feelings*.[11] In this way, reciprocation can be understood as having two senses—a strong and a weak sense—which both fall under the description of a "return of feelings." The first and stronger sense is rather intuitive and refers to the type of reciprocation found in experiences of love. The return of feelings here is a mutual exchange of feelings, or a "requitedness" of feelings: when person A reciprocates in love towards person B, it means that A feels the same way towards B that B feels towards A. The second and weaker sense of reciprocation is the one found in experiences of cravings, where the return of feelings is understood as a *going back to*. Of course, exchanging mutual feelings with another agent and going back to feelings obtained from substance use are different in many ways, but the description of reciprocation as a return of feeling unifies both phenomena while keeping their respective connotations. In other words, the definition and description of reciprocation as a return of feelings allows me to compare the two in a valid and fruitful way.

11 See for example the Oxford English Dictionary's fourth entry of "reciprocate" (https://www.oed.com/dictionary/reciprocate_v?tab=meaning_and_use&tl=true; accessed March 27, 2024) or its fifth entry of "reciprocation" (https://www.oed.com/dictionary/reciprocation_n?tab=meaning_and_use#130644449; accessed March 27, 2024).

Reciprocation in addiction is made possible through two elements. First, the substance one is craving and addicted to is not merely stumbled upon. It might be at first, but once it becomes a drug of choice,[12] it is specifically preferred for some of its qualities. For example, one might prefer marijuana precisely *because* it relaxes and calms them. Second, this chosen addictive substance has the capacity to deliver those qualities for which it is overall preferred, and thus of bringing about a rewarding state. Combined, these two elements give rise to the possibility of reciprocation, understood as a return of feelings in its weak sense. In short, consumption of CO can be seen as an act of reciprocation because it can fulfill a need or a desire which it has itself caused; the consumption of a CO, and the effects it can have on one, are the reasons why a CO is a CO.

The importance of and longing for reciprocation in addiction is made clear through testimonies of addiction. For example, Marc Lewis, a neuroscientist and former addict, recalls the following in his *Memoir of an Addicted Brain* (2011, 44):

> I smoked up a few more times during my remaining days in Toronto. I could sit in my parents' home and feel less than humiliated by the stupidity of our dream, less than angry, no longer sad. I could change the way I felt at will. As long as I had my bag and my papers, I was safe from depression, free of the nausea of shame.

In the last sentence of this passage, it is clear that Lewis gets something from the sole thought of having the substance (marijuana) he wants to consume accessible to him. Knowing that he has the bag and the papers gives him the safety that he needs and seeks, because it offers the possibility of a return, of a "going back" to, feelings. This is a type of non-agential reciprocation: the substance reciprocates because it was chosen to fulfill a need it has led to the creation of, and it can and does fulfill the need when consumed.

Lewis's example is one of many, as it seems that the entire theory of self-medication can be explained though the lens of reciprocation in its weak sense. To summarise, the self-medication hypothesis (SMH) considers that addicted individuals use drugs in order to alleviate and soothe a certain state of being. This hypothesis was first articulated by Edward Khantzian (1985), who has more recently co-authored the book *Understanding Addiction as Self-Medication: Finding Hope Behind the Pain* (2008) along with Mark Albanese. There are two conditions to the SMH. First, drugs are addictive because they have the power to diminish or alleviate certain psychological sufferings. Second, people susceptible to developing an addiction usually have difficulty regulating their emotions, self-esteem, interpersonal relationships, and behaviours. These two aspects combined engender the therapeutic consumption of drugs.

If the SMH is true, it seems like addicted individuals have a longing for reciprocation from their CO that resembles what people in the limerent state feel for their LO. In limerence, one longs for reciprocation, for a return of feelings in the strong sense, from an LO. In cravings, one longs for a return of feelings in the weak sense from their CO, as one hopes to fulfill, through consumption of CO, the need that the CO has itself caused.

c. Dependency of mood on LO's actions, or, more accurately, your interpretation of LO's actions with the probability of reciprocation.

When one is in the state of limerence, one's mood is affected and influenced by every action of one's LO, or rather, by the interpretation of the LO's action and whether or not they indicate a sign or a probability of reciprocation. We can take for example one of Tennov's interviewees' statement (1979, 54–55):

12 See for example O'Connor and Berry (1990) on the topic of drug of choice.

> It was just a little thing. Except it wasn't a little thing. She forgot to wear the pin I had given her, even though I had asked her to be sure to wear it. I wouldn't have minded so much if she had been angry and left it home to get even with me over something. But to forget? There's no way I could have forgotten if she had asked me for anything. It meant I wasn't in her thoughts the way she was in mine. I hoped that she was teasing, that she really was mad at me. Anything but forgetting. Being forgotten was like being dead. I wanted to die. I felt it was all over in that second. I prayed for a sign that she was playing with me. She wasn't. It was ended, but it took me about six more months before I could tolerate believing it.

In this quote, the mood of the person in the limerent state is deeply affected by the possibility of being, or not being, loved back ("It meant I wasn't in her thought the way she was in mine"). Once he realises that the possibility that the love he feels for his LO might be unrequited, he panics and even reveals feeling like wanting to die.

As for the experience of cravings, the probability of reciprocation, which is understood as the possibility of CO to (re)deliver the qualities for which it is craved, also greatly impacts the mood of the craving person.

Lewis tells about one time when he was working late in his university laboratory and discovered morphine bottles in the fridge. Although he had not consumed in a while, on this particular evening, he simply could not control his desire, his craving, for the substance. The thought of the drug being there right next to him, as well as the idea that he would eventually, at some point in the evening, consume some, was especially exciting to him. In this passage, it is clear that the possibility of reciprocation from the morphine really changes his mood: "I'm actually humming this as I get up to replace one rat with the next. I'm humming this and I'm smiling a little to myself, smiling with a sneaky little smile, a sneaky little rat smile. A smile for no one. A smile no one can see. But there is a quickening in my pulse. A part of me has given up" (2011, 205). His facial features cannot hide his pleasure, his excitement, that he knows he will be receiving from injecting morphine. In other words, the possibility of a return of feelings from morphine drastically changes his mood in a positive way.

d. Inability to react limerently to more than one person at a time (exceptions occur only when limerence is at low ebb—early on or in the last fading).

According to Tennov, there will usually only be one LO at a time. This means that, in most cases, a person will be in a state of limerence towards only one person at a time. Sometimes, the person in the limerent state will stop being limerent towards LO_1 because they have started being limerent towards LO_2. In this way, Tennov claims that "the pain of one love seemed only to cease with the advent of the next" (1979, 11). In other words, recovering from feeling in love with LO, perhaps unrequitedly, is achieved through being in love with another, new LO.

The inability to react limerently to more than one person at a time seems similar to what, in the literature and in first-person accounts of addiction, amounts to the concept of "drug of choice." Addicted people will prefer a certain substance (Hopwood et al. 2008) and researchers can even guess what type of substance will be preferred by one depending on one's personality. Aside from personality assessments, drugs of choice are also correlated with the type of suffering one is dealing with and wanting to relieve. This is found to be true through the self-medication hypothesis, which I have previously introduced. In other words, there is a specificity to drug use: in most cases, one uses or seeks to use a particular drug or type of drug, not just any drug. This is not to say that one cannot find many substances pleasurable or desirable, but that an addicted person usually finds one substance in particular as more attractive and as more suitable for his wants and

needs. But in any case, the parallel I want to draw between limerence and cravings specifically concerns the moment when one is craving. And in that moment, there usually is only one substance, only one CO, at a time.

To give an example, there is a passage in Knapp's book where she explains how an alcoholic is always drawn to a specific type of alcohol, and yet this relationship is not exclusive. She names this one drink one's *true love*, and it stands besides secondary loves, past loves, and acquaintances, all represented by different alcoholic drinks. Her personal true love was white wine (1999, 105):

> I was a white wine junkie. Toward the end I'd slug down just about anything but if I had my choice, I'd drink a crisp, cold, dry white, a French Sauvignon Blanc or a Chardonnay from the valleys of northern California. The look of a bottle of white wine in a refrigerator always reassured me somehow, the way it stood there on the shelf, beads of moisture forming on the exterior, the labels forming sharp rectangles of color against the pale golden liquid.

For Knapp, white wine was her drug of choice, even though she mentions that she could have ended up drinking just about anything which contained alcohol. Thus, white wine for Knapp was like an LO to a person in the limerence state: replaceable, but still the preferred option until then.

e. Some fleeting and transient relief from unrequited limerent passion through vivid imagination of action by LO that means reciprocation.

In limerence, it is very common for people to have regular fantasies about their LO through their imagination. Usually, the fantasies will be about LO acting as if reciprocating feelings of love. These mental images will also offer some relief and comfort to the person in the limerent state, who will feel like reciprocation is possible and going to happen in a near future. Take for example Larry's limerence towards Evelyn, which involved a lot of mental imagery (Tennov 1979, 37):

> In my mind I was rehearsing the big moment when I'd suggest we have dinner together and imagine these complex situations that would bring it off. Every time I learned something more about her, I'd incorporate that into my daydream. For example, I first thought of, you know, inviting her to my place to listen to jazz records, and we'd be off and running from there. Then I found out she was a classical musical student and I went into this big idea about getting a pair of super concert tickets through my brother-in-law whose uncle plays the oboe in some orchestra.

In short, in this citation, Larry explains how he often daydreamed about Evelyn and used elements that he knew about her to enhance his fantasies and mental representations of her.

It is very interesting to note that in the experience of cravings as well, mental imagery plays a very important, if not foundational, role. In fact, mental imagery often accompanies cravings and intensifies them.[13] Let's take Natalie's case, found in Marc Lewis's second book, *The Biology of Desire*, as an example to understand better how cravings and mental images are entwined (2015, 52–53):

> She'd be at the restaurant, working her shift, actually enjoying the summer sun streaming through the plate-glass windows, when suddenly it would strike. The thought, the image. It might be triggered

13 According to Andrade et al., "while craving can occur without imagery, it is imagery than provides a rich sensory experience that accentuates relief or pleasure and aggravates feelings of deficit" (2012, 130).

> by smoke rising from the cigarette of someone standing just outside the entrance. Or by the peculiar flattening of a spoon sitting next to the dirty plate on a nearby table, calling to mind the misshapen spoon of the dope-shooter, bent so that the bowl remained horizontal when the spoon lay at rest.

In this example, Natalie's desire to consume heroin starts as a thought, an image, that pops up insidiously, or rather intrusively, in her mind. This seems to be the case of most addicted people who experience desires to consume a certain substance. In fact, according to a study led on intrusive thinking and its suppression by Salkovskis and Reynolds, smokers often have intrusive thoughts about smoking. These thoughts are accompanied by vivid images associated with the act of smoking. As mentioned in the study, the most common mental images were the following (1994, 195): "Yourself smoking cigarettes; cigarette packets; other people smoking, enjoying it; borrowing cigarettes; image of yourself in a relaxed place, enjoying a cigarette; someone offering you a cigarette; image of yourself under stress, having a cigarette to help calm down or cope; yourself buying cigarettes." Moreover, it seems that the intensity of cravings and the vividness of mental images are positively correlated. In fact, Tiffany and Drobes (1990) found that, for smokers, the more vivid were the mental images associated with smoking, the more intense were their desires to smoke. If these particular studies only apply to smoking and related mental imagery, similar findings exist in the case of alcohol. Litt and Cooney (1999) have indeed discovered that alcohol cravings can be induced by asking people to "imagine entering their favorite bar, ordering, holding and tasting a cold, refreshing glass of their favourite beer."[14] In other words, it is sufficient, in order to induce cravings of alcohol, to ask people to form mental images of things and scenes related to alcohol consumption.

The link between mental imagery and cravings can be understood through the desire theory of Intrusive-Elaboration (Kavanagh et al. 2005), which I have previously introduced. Briefly, according to this theory, a desire such as a craving is an intrusive thought that can arise spontaneously in a schematic, verbal, or image format, and which has been cognitively elaborated. Therefore, only elaborated intrusive thoughts become desires. Elaboration consists for example of planning the ways in which the desire will be satisfied, of thinking about what the satisfaction of the desire will lead to. The elaboration usually implies mental imagery and fantasising.

It is interesting to note that, for some people, the mental images in and of themselves are more pleasurable than the consumption of the addictive substance. For example, Andrade et al. (2012, 131) mention that one smoker reported how his imagined cigarette was always the best cigarette ever smoked.

f. Fear of rejection and sometimes incapacitating but always unsettling shyness in LO's presence, especially in the beginning and whenever uncertainty strikes.

In the experience of limerence, one is subjected to an intense fear of rejection from his or her LO. This fear of rejection obviously causes pain, but, as Tennov explains, it also enhances desire because it fuels uncertainty. In other words, limerence is sustained by the hope of reciprocation and its uncertainty. This fear produces many different effects on one in the limerent state, such as shyness and self-doubt. It also has physiological consequences associated with fear responses, such as heart palpitations, flushing, and general weakness. Tennov quotes, as one example of many, Philip's experience in order to justify this component (1979, 49):

> I'd be jumpy out of my head. It was like what you might call stage fright, like going up in front of an audience. My hand would be shaking when I rang the doorbell. When I called her on the phone I felt like I could hear the pulse in my temple louder than the ringing of the phone, and I'd get into such a

14 See Andrade et al. 2012, 128.

> panic listening to the ring and expecting Nelly's voice at the other end that I'd have a moment of relief if no one answered. And when she did answer, I wouldn't know what to say even if I'd gone over the whole thing in my head beforehand. And then whatever I did say never seemed to come out right.

Philip is experiencing intense fear of rejection from his LO, Nelly, which leads him to have the symptoms of what he calls stage fright. Philip's experience is typical of people in the limerent state.

In the case of cravings, one does not fear being rejected *per se* by the substance one is addicted to. However, we can reformulate the fear of rejection identified by Tennov in experiences of limerence by holding that it is actually *a fear of being without*. Perhaps rejection and "withoutness" are not the same thing, but it seems that withoutness can explain the fear of rejection. The idea is to say that rejection is experienced as something negative and frightening because it means that one won't be receiving the love that they want from their LO. In other words, rejection is feared *because* it implies withoutness. Reformulating the fear of rejection in such a way allows me to compare it with the fear of non-availability of the substance—of withoutness—that is so common in addiction.

Knapp, in her memoir, explains how she is afraid of being without the substance, always checking that there is alcohol around for different occasions. As she puts it, what drives the constant checking of the presence of alcohol is the fear of being without it: "There's a dark fear to the feeling of wanting that wine, that vodka, that bourbon: a hungry, abiding fear of being without, being exposed, without your armor" (1999, 58). Jowita Bydlowska says something similar in her own memoir: "And there's fear behind the wanting—the fear that if the wanting gets denied there will be only pain and the fear itself left" (2013, 11). In other words, for both women, fear of being without the substance is an inherent part of their craving experiences.

g. Intensification through adversity.

Tennov cites Joan's experience of being in love in order to show how the uncertainty of being loved back, and even potentially rejected, has a surprising effect on the intensity of limerence: "When I was in love with Barry, I was intensely in love. When I felt he loved me, I was intensely in love and deliriously happy; when he seemed rejecting, I was still intensely in love, only miserable beyond words" (1979, 45). This citation shows how adversity and difficulty in reaching the LO can intensify the feeling of being in love with an LO, as well as increase the degree of involvement in limerence (1979, 46). Doubts and uncertainty are food for limerence. The uncertainty does not, however, have to be real, because it is a matter of perception (1979, 56): if one solely perceives that there is uncertainty in whether or not one's LO feels the same towards one, it is enough to feed the intensity of the limerence.

In cravings, the uncertainty of the achievement of the goal, which would be consumption of the CO, also acts as fuel for the desire. In this interesting passage, Lewis explains how his desire for LSD is, in a way, similar to his desire to be with Lisa. Both are, as he writes, magnified by the uncertainty of their obtainment, which can be understood as a manifestation of adversity (2011, 83):

> As with Lisa, the excited craving of something half-attainable was the most potent elixir. It's true: my attitude toward LSD, a drug, was not much different from my zeal to connect with Lisa, a girl, thanks to a flood of dopamine in my ventral striatum—wanting and wanting and wanting and finally getting, magnified by the uncertainty of the goal.

In other words, the uncertainty of achieving whatever is wanted and craved makes the whole experience of wanting even more intense.

A study led by Dunbar et al. (2000) on smokers also showed how cigarette cravings are higher in intensity during abstinence from smoking than during use. This finding suggests that adversity, which in this case is to be understood as the absence of a fulfilled craving, actually enhances the craving.

h. Acute sensitivity to any act or thought of condition that can be interpreted favorably, and an extraordinary ability to devise or invent "reasonable" explanations for how the neutrality that the disinterested observer might see is in fact a sign of hidden passion in the LO.

The component of acute sensitivity in the experience of being in love means to point out how in limerence, any action from an LO will, as much as possible, be interpreted as a positive sign of reciprocation. In her book, Tennov gives the example of Dr. Vesteroy, whose LO at the time was a female colleague. In his testimony, it is very clear how the LO's actions are always interpreted in a way that could indicate reciprocation (1979, 59):

> At first, I'd set up little tests. I'd say that if at the next meeting she elects to sit beside me or facing me, I will count it as proof that this madness is not unilateral. But when she chose a seat farthest from me, or one which made it very difficult for us to look at one another, I realized that the test was not a test at all. No matter what she did, I could interpret it in my favor.

Even when his own tests would fail, Dr. Vesteroy tried to explain his LO's behaviour in a way that did not exclude the possibility of reciprocation from his LO.

In the case of cravings, an addicted person who is craving a certain substance will most likely interpret facts and feelings in a way that makes it likely that the substance and its consumption will bring about the desired rewarding state for which it is craved. In other words, the drug is perceived in favour of its possible reciprocation.

For example, Bill Clegg writes in his addiction memoir how smoking crack helped him deal with the harsh reality that it had itself brought out (2010, 34):

> If anyone had stopped to watch me go to the cash machine and withdraw stacks of bills, several times because of the $200 transaction limit, then head out to an idling van with tinted windows, and return minutes later with bulging pockets, it wouldn't take much imagination to understand what had just transpired. As obvious and sloppy as I know the whole operation is, I know that once I get back to the room and take a big hit off one of the crystal-clear new stems, everything will be okay. That all the grim and alarming truths barking loudly around me will vanish in a blast of smoke.

Moreover, in an earlier passage (2010, 32), he explains how even though he missed two flights and stood up his boyfriend because of his smoking frenzy, he could not have been happier because he had consumed and was feeling high. Clegg's case clearly shows how, in addiction, when one is craving a certain drug, one will interpret and perceive things in a way that will favour its consumption. It is as if nothing else mattered other than using, and thus reciprocation, taken as the possibility of a return of feelings through consumption of the CO.

i. An aching of the "heart" when the uncertainty is strong.

Tennov explains that limerence is located in the heart, rather than in the arms, as when one is feeling affection, or in the genital area as in the feeling of a sexual desire (1979, 64). Therefore, for Tennov, there is a distinction between limerence and other feelings relative to love through the body parts they are associated with. Although this element is interesting because it draws from bodily feelings in order to make sense of an emotional state, it seems like it is not convincing for at least two reasons.

First, purely physiological sensations can be confused with emotions such as love. Karen Jones (2008) presents a case of a woman who believes she is in love with someone because she feels tingly in her stomach. However, it turns out to be a stomach ulcer instead of love. Second, it is not scientifically clear that, if there were to be a physical body part involved in love, it would be the heart specifically. Although it is true that emotions are associated with specific bodily patterns and reactions, a study lead by Nummenmaa et al. (2013) on the bodily maps of emotions shows how an emotion like love is felt throughout the entire body, and not just in the heart. Therefore, because of the lack of empirical data supporting this element, I do not think it should be considered in our comparison between limerence and the experience of cravings.

j. Buoyancy when reciprocation seems evident.

When things are going well for one in the limerent state, such as when reciprocation is thought to take place, one experiences feelings of buoyancy, of floating or of walking on air. Recall for example the way that Joan describes her feelings for Barry: "When I was intensely in love with Barry, I was intensely in love. When I felt he loved me, I was intensely in love and deliriously happy; when he seemed rejecting, I was still intensely in love, only miserable beyond words" (1979, 45). She mentions precisely that when she felt like there was some type of reciprocation, she was, in return, *deliriously* happy.

This feeling of buoyancy is also present in the craving experience for an addicted individual. As a matter of fact, when things are going well and the craving is therefore thought to be satisfiable, the craving individual can feel ecstatic, motivated, and strong. Take for example Bydlowska, who explains in her memoir how the desire driving her to drink made her feel thrilled (2013, 39): "I don't consciously think of dying when I'm drinking. In fact, the desire that I imagine drives my drinking is the desire to live, to live loudly and freely, without any care. I want to jump, want to run, want to want!" Lewis also has a passage where he describes the euphoria he feels when he knows that he will soon be able to consume his drug of choice, heroin (2011, 129): "Elation surged through me: it was really going to happen. Today. Soon." Therefore, if we understand reciprocation as the return of feelings in its weak sense, it is clear that the indication of such reciprocation in the eyes of addicted individuals leads to a feeling of glee. Just like someone in love, someone craving a substance is affected by the thought of reciprocation.

k. A general intensity of feeling that leaves other concerns in the background.

Without a doubt, limerence is considered to be a very intense and overwhelming mental state that affects attention and perception. In fact, when one is limerent, Tennov explains, one is in a condition of "sustained alertness, a heightening of awareness, and an enormous fund of energy to deploy in pursuit of the limerent aim" (1979, 62). Here is a quote from the prologue of her book that illustrates well how one's attention is modified when in the limerent state (1979, xiv):

> Everything reminds me of you. I try to read, but four times on a single page some word begins the lightning chain of associations that summons my mind away from my work, and I must struggle to

> return my attention to the task at hand. Often I give up easily, leave my desk, and throw myself down on my bed, where my body lies still while my imagination constructs long and involved and plausible reasons to believe that you love me.

In this quote, the person in the limerent state is really struggling to remain focused on tasks unrelated to an LO. This is because the intensity of the limerence consumes thoughts and attention processes.

As for the role of cravings on attention and perception, there are numerous examples, either coming from personal stories of addiction or large-scale studies, that show how drugs affect the brain in its mechanisms of perception and attention. To give a first example of a testimony, take this quote from Lewis's book on his experience with addiction: "By ten I say goodnight and shut my door. I haven't shot drugs in a long time, and the anticipation is a fist grabbing me by the collar and pulling me forward through time" (2011, 157). In this passage, Lewis literally writes that the anticipation of using drugs is pulling him forward through time. His perception of time is therefore distorted by the intensity of his craving and by his knowledge that he will soon satisfy his craving. Indeed, his perception is modified by the knowledge that the drug he craves will reciprocate.

There is, moreover, a neuroscientific theory that can explain the effect of wanting on perception and attention. Robinson and Berridge have, in fact, developed the incentive salience theory of addiction, which is now well-known and recognised in the literature on addiction. This theory stipulates that "the defining characteristics of addiction (craving and relapse) are due directly to drug-induced changes in those functions normally subserved by a neural system that undergoes sensitization-related neuroadaptations" (1993, 249). In other words, according to the authors, neuroadaptations in the brain occur because of the chronic consumption of drugs, which can explain why addiction manifests itself through frequent relapses and cravings. The neural system in question controls a precise psychological function, which is called salience attribution in perception and mental representation of stimuli. But as this system is changed through neuroadaptations, due to the consumption of drugs, it becomes hypersensitive in regards to certain stimuli associated with consumption (sights, smells, or any other detail that can be associated with the drug). The salience of these specific stimuli associated with consumption and drugs makes it very hard to control consumption in part by triggering cravings. In short, to go back to the way cravings affect attention and perception, the incentive salience theory helps us to understand that when one is craving a certain substance, one is pulled, in perception and attention, towards stimuli associated with the CO because of the neuroadaptations that have taken place in the salience attribution system. Thus, without a doubt, just like one's attention is significantly altered by limerence, attention and perception are influenced by cravings.

l. A remarkable ability to emphasise what is truly admirable in LO and to avoid dwelling on the negative, even to respond with a compassion for the negative and to render it, emotionally if not perceptually, into another positive attribute.

When one is limerent, everything about one's LO is considered to be a positive trait, and when there is a risk of perceiving something that could be interpreted as a negative trait, or as a less positive one, that risk is avoided. Fred's account in Tennov's book is particularly revealing of this aspect (1979, 98):

> It seems to me that being romantically attracted to Laura means that I am bending my image of her until it is distorted. Things that might produce an unpleasant picture, I simply do not see. When she appears by relatively objective standards, beautiful and capable, I look long and hard. But when she is not at her best, when I catch her face in an unflattering angle, I turn my eyes away.

Fred prefers to look away when he might catch Laura, his LO, in an unfavourable light, because he is so in love with her. In a later passage of his diary, Fred explains how Laura shows no interest in his work in architecture, which points towards the fact that she would not be a viable girlfriend for him to have, but he nevertheless still holds on to his feelings for her and remains in a state of limerence. He even tries to convince himself of the advantage of Laura being uninterested in his work.

In addiction, there are many examples of addicted persons who, while craving to consume a certain substance, suddenly only perceive the good or inoffensive traits of their CO. Take for example this excerpt from Nic Sheff's autobiography on addiction, where he recalls the moment when he is headed to a friend's place to consume some marijuana after a bout of abstinence (2011, 170):

> She leads me down a couple more blocks, and I talk pretty incessantly the whole time, even though my mind is somewhere else entirely. I mean, basically I'm just going over why this is all okay—over and fucking over again. 'Cause, see, the thing is, the reason I got addicted in the first place was because the drugs took my terror and depression away. But now I've finally learned how to love and value myself. I've grown and changed. So there's no reason why drinking or smoking pot should be a problem.

In this passage, Nic explains how he is convincing himself of the harmlessness of drinking or smoking, which is what he is about to do, and more importantly, what he desires to do. In that moment, he is therefore interpreting his attitude and actions in ways that make drug use favourable ("I've finally learned how to love and value myself"). Indeed, and as he puts it, he has grown and changed from his past as a former addict, so he no longer needs to worry about consuming substances like pot or alcohol. In our view, the resemblance between Fred's way of coping with his perception and interpretation of his LO, and of Nic's way of coping with his CO, is, although somewhat uncanny, conspicuous.

5. Conclusion

In this article, I have tried to show how the experience of cravings is like the experience of being in love with someone. Although many comparisons have been made between addiction and love, none have tried to show that it is the processes in addiction which are similar to the ones in love. Indeed, in the past, researchers have taken the opposite route in order to prove how people in love act and react as if they were addicted. Yet addicted individuals themselves use the vocabulary of love and romance in order to describe how it feels to be addicted. Therefore, it seems warranted to think that addiction resembles love. More precisely, and as I have argued in this article, there are reasons to think that some part of the experience of addiction (cravings) is like some part of the experience of love (being in love). Furthermore, this comparison between the two phenomena has allowed me to introduce the notion of reciprocation in the addiction literature. I define reciprocation as the possibility of a return of feelings, where in love it holds a rather strong sense (a mutual exchange of feelings of love) and in addiction a weaker one (a "going back to" feelings). Reciprocation in addiction is enabled through two features: first, a substance is chosen for its rewarding effects and properties; second, a substance, through consumption, can deliver the effects it has been chosen and is craved for, and can therefore keep being rewarding. This concept of reciprocation applied to addiction seems noteworthy because it highlights the fact that a substance can repeatedly satisfy one through consumption, and that it is precisely because of this possibility to repeat satisfaction through consumption that it is craved. A craved object is craved because it has proven to be rewarding for one *and* because it can continue being rewarding. Thus, reciprocation appears timeless, only acquiring temporal properties through the locality

of an act of consumption. Maybe this can help to explain why one finds it so hard to quit: overcoming the longing for reciprocation requires that one overcome the timelessness of the longing for reciprocation. Finally, thinking of the experience of cravings as resembling the experience of being in love accentuates the fact that addiction is, at least in one of its senses, a way of seeing and interpreting the world. Just like one in love is interpreting life's events through the lens of being in love, one who craves an addictive substance is interpreting life's events through the lens of being addicted.

References

Andrade, J., J. May, and D. Kavanagh. 2012. "Sensory Imagery in Craving: From Cognitive Psychology to New Treatments for Addiction." *Journal of Experimental Psychopathology* 3 (2): 127–45.

Bruehl, A. M., D. H. Lende, M. Schwartz, C. E. Sterk, and K. Elifson. 2006. "Craving and Control: Methamphetamine Users' Narratives." *Journal of Psychoactive Drugs* 38 (Sup. 3): 385–92.

Bydlowska, J. 2013. *Drunk Mom: A Memoir*. Doubleday.

Clegg, B. 2010. *Portrait of an Addict as a Young Man: A Memoir*. Hachette.

Dhawan, A., R. Kumar, S. Yadav, and B. M Tripathi. 2002. "The Enigma of Craving." *Indian Journal of Psychiatry* 44 (2): 138–43.

Dill, B. and R. Holton. 2014. "The Addict in Us All." *Frontiers in Psychiatry* 5 (139).

Dunbar, M. S., D. Scharf, T. Kirchner, and S. Shiffman. 2010. "Do Smokers Crave Cigarettes in Some Smoking Situations More than Others? Situational Correlates of Craving when Smoking." *Nicotine & Tobacco Research* 12 (3): 226–34.

Earp, B. D., O. A. Wudarczyk, B. Foddy, and J. Savulescu. 2017. "Addicted to Love: What Is Love Addiction and When Should It Be Treated?" *Philosophy, Psychiatry, & Psychology* 24 (1): 77–92.

Elster, J. 1999. *Strong Feelings: Emotion, Addiction, and Human Behavior*. MIT Press.

Fennell, E., dir. 2023. *Saltburn*. UK: Amazon MGM Studios & Warner Bros. Pictures, 2023.

Fisher, H. E., X. Xu, A. Aron, and L. L. Brown. 2016. "Intense, Passionate, Romantic Love: A Natural Addiction? How the Fields that Investigate Romance and Substance Abuse Can Inform Each Other." *Frontiers in Psychology* 7: 687.

Foddy, B. and J. Savulescu. 2007. "Addiction Is Not an Affliction: Addictive Desires Are Merely Pleasure-Oriented Desires." *The American Journal of Bioethics* 7 (1): 29–32.

Grisel, J. 2020. *Never Enough: The Neuroscience and Experience of Addiction*. Anchor.

Henden, E. 2008. "What Is Self-Control?" *Philosophical Psychology* 21 (1): 69–90.

Hopwood, C., K. L. Baker, and L. C. Morey. 2008. "Personality and Drugs of Choice." *Personality and Individual Differences* 44 (6): 1413–21.

Hormes, J. M. and P. Rozin. 2010. "Does 'Craving' Carve Nature at the Joints? Absence of a Synonym for Craving in Many Languages." *Addictive Behaviors* 35 (5): 459–63.

Jones, K. 2008. "How to Change the Past." In *Practical Identity and Narrative Agency*, edited by K. Atkins and C. Mackenzie, 269–88. Routledge.

Kavanagh, D. J., J. Andrade, and J. May. 2005. "Imaginary Relish and Exquisite Torture: The Elaborated Intrusion Theory of Desire." *Psychological Review* 112 (2): 446.

Khantzian, E. J. 1985. "The Self-Medication Hypothesis of Addictive Disorders: Focus on Heroin and Cocaine Dependence." *The American Journal of Psychiatry* 142 (11): 1259–64.

Khantzian, E. J. and M. J. Albanese. 2008. *Understanding Addiction as Self-Medication*. Rowman & Littlefield.

Knapp, C. 1999. *Drinking: A Love Story*. Dial Press.

Kozlowski, L. T. and D. A. Wilkinson. 1987. "Use and Misuse of the Concept of Craving by Alcohol, Tobacco, and Drug Researchers." *British Journal of Addiction* 82 (1): 31–36.

Lamm, H. and U. Wiesmann. 1997. "Subjective Attributes of Attraction: How People Characterize Their Liking, Their Love, and Their Being in Love." *Personal Relationships* 4 (3): 271–84.

Lewis, M. 2011. *Memoirs of an Addicted Brain: A Neuroscientist Examines His Former Life on Drugs*. PublicAffairs.

———. 2015. *The Biology of Desire: Why Addiction Is Not a Disease*. PublicAffairs.

Litt, M. D. and N. L. Cooney. 1999. "Inducing Craving for Alcohol in the Laboratory." *Alcohol Research & Health* 23 (3): 174.

Loewenstein, G. 1999. "A Visceral Account of Addiction." In *Getting Hooked: Rationality and Addiction*, edited by J. Elster and O. J. Skog, 235–64. Cambridge University Press.

Lopez-Cantero, P. 2022. "Falling in Love." In *Love: Past, Present and Future*, edited by N. McKeever, J. Saunders, and A. Grahlé, 107–23. Routledge.

Marlatt, G. A. 1978. “Craving for Alcohol, Loss of Control, and Relapse: A Cognitive-Behavioral Analysis.” In *Alcoholism: New Directions in Behavioral Research and Treatment*, edited by P. E. Nathan, G. A. Marlatt, and T. Løberg, 271–314. Plenum Press.

Nozick, R. 1989. *The Examined Life: Philosophical Meditations*. Simon & Schuster.

Nummenmaaa, L., E. Glereana, R. Harib, and J. K. Hietanen. 2013. “Bodily Maps of Emotions.” *PNAS* 111 (2): 646–51.

O’Connor, L. and J. W. Berry. 1990. “The Drug-of-Choice Phenomenon: Why Addicts Start Using Their Preferred Drug.” *Journal of Psychoactive Drugs* 22 (3): 305–11.

Ortega y Gasset, J. 1957. *On Love: Aspects of a Single Theme*. World Publishing Company.

Peele, S. and A. Brodsky. 1975. *Love and Addiction*. Taplinger.

Reynaud, M., L. Karila, L. Blecha, and A. Benyamina. 2010. “Is Love Passion an Addictive Disorder?” *The American Journal of Drug and Alcohol Abuse*, *36*(5): 261–7.

Robinson, T. E. and K. C. Berridge. 1993. “The Neural Basis of Drug Craving: An Incentive-Sensitization Theory of Addiction.” *Brain Research Reviews* 18 (3): 247–91.

Salkovskis, P. M. and M. Reynolds. 1994. “Thought Suppression and Smoking Cessation.” *Behaviour Research and Therapy* 32 (2): 193–201.

Sheff, N. 2011. *We All Fall Down: Living with Addiction*. Little, Brown Books for Young Readers.

Sripada, C. 2018. “Addiction and Fallibility”. *Journal of Philosophy* 115 (11): 569–587.

Sussman, S. 2010. “Love Addiction: Definition, Etiology, Treatment.” *Sexual Addiction & Compulsivity* 17 (1): 31–45.

Tennov, D. 1979. *Love and Limerence: The Experience of Being in Love*. Scarborough House.

Tiffany, S. T. and D. J. Drobes. 1990. “Imagery and Smoking Urges: The Manipulation of Affective Content.” *Addictive Behaviors* 15 (6): 531–9.

Timmreck, T. C. 1990. “Overcoming the Loss of a Love: Preventing Love Addiction and Promoting Positive Emotional Health.” *Psychological Reports* 66 (2): 515–28.

Tracy, J. I. 1994. “Assessing the Relationship between Craving and Relapse.” *Drug and Alcohol Review* 13 (1): 71–7.

Weingarten, H. P. and D. Elston. 1990. “The Phenomenology of Food Cravings.” *Appetite* 15 (3): 231–46.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Received: 6 February 2024
Accepted: 6 June 2024
Published: 30 June 2024

Loving a Narrator

Pilar Lopez-Cantero – Tilburg University, The Netherlands

Abstract

We love people because of who they are, but can the idea of "who they are" be explained through a property that everyone has, such as agency? David Velleman believes this to be the case, and argues that love is an appraisal of a person's incomparable value, which disarms the lover's emotional defences. Modelling love on Kantian respect, Velleman claims that love is a response to a person's rational nature, indirectly perceived though her empirical persona—her observable traits and behaviours, which are imperfect representations of the value of rational nature. An important problem for Velleman's account is that it seems incompatible with two widely shared assumptions about love: that love is personal (so it cannot be analogous to impersonal respect) and that love is selective (so it cannot be based on a property shared by all individuals). Here, I propose a re-formulation of Velleman's view that avoids those objections while preserving the idea that love is an evaluation of the loved person's agency. Specifically, love is an evaluation of a person's inner narrator, which is also a person's capacity to act for reasons—to make actions intelligible to herself. The inner narrator is perceived through that person's observable narrative: her actions and interpretations in interaction, which are a product, and not an imperfect representation, of the inner narrator. In the re-formulated view, love is both personal and selective, but still an agential process that involves both the lover's and the loved person's capacities to make sense of the world.

Keywords: love, relationships, narrative, agency, moral emotions

In "Love as a Moral Emotion" (1999), David Velleman offers a view on love for persons that is both unique and controversial. For Velleman, love is a valuing response that disarms the lover's emotional defences, and, because it is directed at a person's rational nature, love is not at odds with morality. The view is unique because, unlike other cognitivists, Velleman focuses on the phenomenology and the morality of love—two aspects that at the time had been neglected by philosophers who described love as an evaluation and who focused mostly on rationality. The view is controversial because Velleman argues that love is an evaluation of the loved person's true self. The self is understood within a Kantian framework as that person's rational nature, and so love is directed at the same object as respect. This view is, then, a radical departure from mainstream cognitivist views, which see love as an evaluation of the loved person's observable qualities.

One particularly damning line of objection is that, under Velleman's account, love losses its personal and selective character. The objection is twofold. First, if both love and respect are modes of appreciation of the same object, it is unclear how we can differentiate between personal love and impersonal respect. Second, if love is an appreciation of rational nature rather than the observable qualities of the loved person, it is unclear why we love some people and not others, given that everyone has rational nature—people seem therefore to be interchangeable as objects of love.

Velleman pre-empts this objection and tries to accommodate the personal and selective character of love by distinguishing between the final object of love's evaluation—the loved person's rational nature—and the immediate object of love's evaluation: the loved person's observable qualities, which Velleman calls her "empirical persona." As I show, this answer does not succeed. But that does not make the view unsalvageable; in fact, I argue that it is possible to avoid the objection, drawing from Velleman's own writings. I propose a re-formulation of Velleman's view that substitutes the Kantian framework with Velleman's views on narrative agency and narrative explanation.

The argument goes as follows. In Section 1, I situate and summarise Velleman's view on love and briefly motivate my aim. In Section 2, I present the objections based on the indistinguishability of love and respect and the incompatibility of the view with selective love, as developed by Edward Harcourt (2009) and Benjamin Bagley (2015). I argue that Velleman's available responses are incompatible with each other, but I present the potential alternative of re-interpreting the view while maintaining its structure. In Section 3, I draw from Velleman's work on narrative theory to argue for a replacement of rational nature with inner narrator, and empirical persona with observable narrative. This proposal, I argue, avoids both objections. In Section 4, I assess the re-formulated view against objections based on the commitment to narrative theory and the role of love in moral life.

1. Situating Velleman's View

To illustrate the significance of Velleman's account, let us start with a brief look into contemporary philosophy of love. For most part of the 20th century, many philosophers of love could be placed on one side of what Bennett Helm (2010, 10) has called the cognitive–conative divide. On the one side, cognitive views see love as a way of appreciating the person who is loved; a mode of valuing that person. On the other side, conative views do not see love as a mode of valuing, but as a type of concern that causes or entails the lover's bestowal of value upon the loved person.

Often, philosophers on either side have disagreed on what has become known as the matter of the reasons of love. It is a rather murky debate, and it is not always clear exactly which question is being asked. At first glance, it seems that the question being asked is: "What is love? (a mode of valuing or a mode of concern?)" But this question has quickly turned into another one: "Do we love for reasons?" This version of the question has in turn been understood in two seemingly similar but distinct ways. Some authors are concerned with the question "Is love the sort of attitude that can be justified in terms of reasons?" Cognitivists answer this question in the affirmative (Naar 2017), whereas those favouring conative views answer the question in the negative (Frankfurt 2004; Zangwill 2013). Other authors ask a different question under the guise of the one above: "What are, if any, *good* reasons for love?" which is often taken to be equivalent to the slightly different question: "What reasons, if any, justify love *for this person and not another*?" This latter question has come to be known as the question of selectivity, often posed by cognitivists as a question about the normative force of selective reasons: "Why *should* you love this person and not another?" That is, if we select some people and not others in our loving evaluation, what is the (rational and/or moral) justification for that selection, given that all people are in principle equally lovable? The relationship-based answer to this version of the question is that the source of justifying reasons for love is shared history (Kolodny 2003), while quality-based views point at the loved person's traits or observable qualities—such as her personality, her virtues, or her hair colour (Keller 2000; Abramson and Leite 2011).

The progression described above shows how philosophy of love moved from a descriptive question—"What *is* love?"—to a normative question—"Why *should* you love this person and not another?"—often within the same argument. Along the way, philosophy of love became paradoxically bereft of emotion. The focus on normativity and justification left little to no space for discussing how joyful, fearsome, and soul-destroying loving another person can be. In that way, the phenomenological component of love remained largely ignored by analytic philosophers from both sides of the divide, often embroiled in a discussion about rationality and direction of fit.

Velleman's account is to be situated at the height of that dialectic, when the stark division between cognitivist and conative views and the focus on justification was seldom questioned. It is against this background that "Love as a Moral Emotion" emerges as a notable exception, insofar as Velleman presents a cognitivist view that answers the descriptive question ("What is love?") while putting the phenomenological element of love at the centre of his answer to that question. For Velleman, love is an appreciation of the incomparable value of the loved person, which motivates the lover to arrest her emotional defences or "suspend [her] emotional protection" (1999, 362). Velleman does not concern himself with the rationality of love, or the normative force of reasons for love. Instead, he sets out to offer a description of what love is—an evaluation—with a distinct double aim: to argue that love is an essential component of moral life, and to show that a moralistic view of love can offer a satisfactory answer to the question of selectivity.

Let us start with the latter aim. According to Velleman, most answers to the question of selectivity wrongly focus on observable qualities of the loved person, such as beauty or intelligence. However, we do not want to be loved for qualities that may wear off with time, or that a number of other people may have: we want to be loved because we are special, unique, and irreplaceable (363). "We are like the girl who wants to be loved not for her yellow hair—and not, we should add, for her mind or her sense of humour, either—because she wants to be loved, as she puts it, 'for myself alone'" (363), he says in a reference to Yeats's poem.[1] We want to be loved for our true selves, which arguably do not change when our hair goes grey, when we start losing our memory due to Alzheimer's, or when our jokes are not funny anymore. Up until this point, his account seems to correspond with some plausible folk intuitions on our desires about love.

This correspondence with folk intuitions ends with Velleman's view of true selves, which are the object of evaluation in love, as the loved person's rational nature, understood in the Kantian sense of "a capacity to be actuated by reasons" (365).2 According to Velleman's interpretation of Kant, instances of rational nature or 'self-existent ends' (that is, persons) require a specific evaluative response from others: respect. Not responding to rational nature with respect is both a rational and moral failure. To appreciate oneself as a self-existent end, one needs to respect other self-existent ends: an instance of rational nature "cannot first take itself seriously if it treats instances of itself as nothing more than mere means to an end" (366). However, respect is not the only way of appreciating the value of persons: we can also appreciate the value of persons by loving them.

1 "But I can get a hair-dye/ And set such colour there,/ Brown, or black, or carrot,/ That young men in despair/ May love me for myself alone/ And not my yellow hair." ("For Anne Gregory"; Yeats 1956).

2 Velleman (365) makes this equivalence explicit in the following passage: "Recall the following tenets of Kantian theory: that the rational nature whose value commands respect is the capacity to be actuated by reasons; that the capacity to be actuated by reasons is also the capacity to have a good will; and that the capacity for a rational and consequently good will is that better side of a person which constitutes his true self."

Velleman says that love and respect are different "modes of valuation" (366) or ways of appreciating the same object: a person's rational nature. Love and respect are distinct in two ways. Firstly, respect is a required response to the value of a person's rational nature, while love is a possible, but not required, response to the value of a person's rational nature. In turn, rational nature is itself both a capacity for respect (as seen above) and a capacity to love (365). Secondly, love and respect motivate different reactions: while respect arrests our tendency to treat other people as means, love arrests our tendency to be emotionally defensive towards others (362).

Velleman acknowledges that, at face value, his proposal does not meet the condition of selectivity. After all, he argues that love does not discriminate between people. He is also aware of the counterintuitiveness of the claim that rational natures are what we love about people, since we do not directly perceive rational natures. What we do perceive are the observable qualities of the loved person that other cognitivists identify as reasons for love: the loved person's personality, her mannerisms, even "the way he wears his hat and sips his tea … or the way he walks and the way he talks" (371). Velleman calls these observable qualities of the loved person her "empirical persona": "the manifest person, embodied in flesh and blood and accessible to the senses" (371). Thus, Velleman's complete answer to the descriptive question of what love is, including his answer to the question of selectivity, goes as follows: love is an awareness of the incomparable value of the loved person's rational nature, as represented by her observable empirical persona, and which results in the lowering of the lover's emotional defences.

Now let us turn to the aim of showing that love is a moral emotion, which is in fact Velleman's primary aim. Besides the debate described above among philosophers of love, Velleman's view is situated in another long-standing discussion among moral philosophers about the apparent incompatibility between love and moral choice: given that moral choice should not discriminate between people, and love *is* a discrimination between people, it seems that love is incomparable with morality.[3] Velleman's answer to the challenge is to argue that it is based on a false assumption. In his view, love does *not* involve a discrimination between people, since loving a person is to value them as having dignity, which itself precludes comparison: "[B]eing treasured as special doesn't entail being compared favorably with others; it rather entails being seen to have a value that forbids comparison" (370). Since love does not discriminate, and it is a mode of valuing people, love is not at odds with morality: in fact, love is an essential component of moral life.

Velleman's account has generated a great deal of disagreement. Here I focus on a specific objection raised by Harcourt (2009) and Bagley (2015) on the apparent incompatibility of the account with the personal and selective features of love—the fact that we love some people and not others based on something specific about those people. Before proceeding, three brief clarifications are needed. The first is that my re-formulation of Velleman's view is posited as a solution to the objection based on the personal and selective character of love only. I take as a starting assumption that love for persons is selective and not impersonal, and so I focus on answering the objection to that starting assumption, if Velleman's view is to be considered at all.[4]

3 Sophie Grace Chappell (2014, 80–84) offers a useful summary of the challenge and the different responses to it—like Velleman, she disputes the formulation of the challenge itself.

4 I leave aside objections formulated by Kennett (2008), Abramson and Leite (2011), Setya (2014), or Foster (2018); I also leave out an objection from Harcourt (2009) on the relation between dignity and value, and a myriad of rejections of Velleman's view in discussions about reasons for love. Also, in this paper I am not questioning Velleman's wider interpretation of Kant, which has been qualified as "revisionist" (Harcourt 2009, 349). These objections may or may not affect the re-formulated view, but that is beyond the scope of this paper.

The second clarification is that, in re-formulating Velleman's view, I am not necessarily endorsing the view myself, for two reasons. The first is that I am not convinced that the claim that love is a mode of evaluation is correct, or if instead the correct claim is that love requires or merely involves evaluation, or—my preferred option—that love depends on or is grounded in evaluation. The second is that Velleman's views on agency and practical reason are not always clear or consistent, so I acknowledge that it is possible to disagree with my interpretation of Velleman. I will offer my reasoning for understanding Velleman in the way that I do, but I will not attempt to prove that my interpretation of Velleman is the correct one. For these reasons, my claim here is conditional and qualified: *if* love is a mode of valuing, the view that what is valued is the loved person's agency is a *plausible* view, and such a view can be obtained from *a reasonable interpretation* of Velleman's views on agency and practical reason.

My argumentative stance leads to the third clarificatory question: if I am not endorsing Velleman's view, why engage with it at all? In recent years, there have been several proposals that overcome the obstacles of the cognitive–conative divide and the exaggerated focus on rationality.[5] Why not simply abandon cognitivism and move on? One reason to preserve Velleman's view is the historical and current significance of the account, which is cited almost without exception in any review of contemporary philosophy of love, and which is widely regarded as a major contribution to the debate. Philosophical relevance is a good enough reason to continue engaging with Velleman's view on love, but there is a better, more concrete reason to do so.

The re-formulated version of Velleman's account can be a useful addition to the discussion on the role of agency in love. According to agency-based views, love is something that we *do* and thus is an expression of our agency, and not merely an experience that we undergo passively (Kühler 2014; Schaubroeck 2014; Ebels-Duggan 2008; Candiotto and De Jaegher 2021; Spreeuwenberg 2021). These views are mostly focused on the agency of lovers and the interactions between them; the people in the equation that *do* the loving, which in virtue of the focus on activity, are all the people involved. To put it coarsely, there is no more "lover" and "loved person" in these discussions. To put it more precisely, since these accounts are focused on action and interaction, there is less focus on the loved person as an object of evaluation, which makes sense given that "being an object of evaluation" seems to be a passive feature of persons, and thus in direct opposition to their roles as agents. However, the fact that these authors focus on loving instead of being loved does not mean that they refuse to engage with the latter issue—explicit calls to integrate the "passive" element of love can be found in Michael Kühler's (2014) and Kyla Ebels-Duggan's (2019) proposals, for example. Re-formulating Velleman's view so that there is a plausible view of agency of the loved person as the object of love is a valuable addition to the discussion of love and agency in that sense. As I will later argue, the re-formulated view also accounts for the role of the agency of the lover in love's evaluation. However, these two aims—showing that agency can be the object of love and that the agency of the lover is involved in love's evaluation—are unachievable without resolving the objections about the personal and selective character of love first.

5 For example, by proposing a double direction of fit that acknowledges both the cognitive and the conative aspects of love (Helm 2010; Bagley 2015; Han 2021); by arguing that both rational and irrational love can exist (Kroeker 2019; Saunders and McKeever 2022); or by differentiating subjective reasons for love from objective normative reasons (Jollimore 2011; Schaubroeck 2014; Jeske 2017; Naar 2017).

2. Personal and Selective Love

As seen above, in Velleman's account love has a final object of evaluation—rational nature—and an immediate object of evaluation—an empirical persona. Introducing the empirical persona as the immediate object of evaluation is not an argumentative manoeuvre on Velleman's part to get out of the acknowledged counterintuitiveness of his view. Recall that respect is a required response to others in order to appreciate oneself as a self-existent end. Although love is an optional response to others, it is nevertheless necessary in order to lower one's emotional defences. Only by loving can we do this. And we can only love by engaging with an empirical persona:

> The manifest person is the one against whom we have emotional defenses, and he must disarm them, if he can, with his manifest qualities. ... [U]nless we actually *see* a person in the human being confronting us, we won't be moved to love; and we can see the person only by seeing him in or through his empirical persona. (1999, 371)

The selective character of love—that we love some people and not others—is a result of practical limitations imposed by the cognitive demandingness of love. According to Velleman (372) we are "constitutionally limited" in our capacity to love, given that loving some people takes up all of the capacity for attention we have. More importantly, we are not well equipped to perceive a person's value through her empirical persona, or to represent our own value through our empirical personae:

> The human body and human behavior are imperfect expressions of personhood, and we are imperfect interpreters. Hence the value that makes someone eligible to be loved does not necessarily make him lovable in our eyes. Whether someone is lovable depends on how well his value as a person is expressed or symbolized by us by his empirical persona. Someone's persona may not speak very clearly of his value as a person, or may not speak in ways that are clear to us. (372)

Velleman's solution for selectivity is unsuccessful, as illustrated by objections put forward by Harcourt (2009) and Bagley (2015) on the basis of two distinct but related aspects of love: the fact that love is personal and the fact that the loved person should not be interchangeable with another.

The first objection is that Velleman's account seems to do away with personal love. Personal love is not just a generic love for persons, understood as agapeic or universal love for humanity in general, but love for particular individuals. Harcourt (2009, 3) argues that, "[i]f we love only a subset of the people we respect because we miss their eligibility to be loved, the fact that we do not love more people is our mistake. But if it is a mistake, we must be able to conceive of correcting it." Similarly, Bagley (2015, 485) says that, on Velleman's account, love is "a genuine rational imperfection, one we have reason to work to overcome." Both Harcourt (2009, 3) and Bagley (2015, 485) conclude that, given that love is a suboptimal response on Velleman's account, then his view can only possibly explain impersonal love.

The second objection is that pointing at rational natures as the final objects of love's evaluation seems to establish a requirement for loving every person, since everyone has rational nature. This entails that people are interchangeable as objects of love: "If love is elicited by a property that all rational beings share, it seems to follow that it cannot matter which rational being one loves" (Harcourt 2009, 350). Bagley (2015, 485) says that Velleman misses the phenomenological significance that selectivity has for the lover, who needs to experience her love for a particular individual as "more than merely optional." This significance is not captured in

Velleman's selectivity, understood in terms of psychological limitations, which according to Bagley is a trivial or superficial kind of selectivity. Instead, in Velleman's view, the fact that we love one person and not another is "a strictly causal matter, just an incidental quirk of [our] psychology" (484).

Velleman anticipates the first objection and explicitly rejects it: "In merely respecting rather than loving [someone], we do not assess them as lower in value" (1999, 372). That is, he explicitly says that respect does not stand higher than love in a hierarchy of modes of valuing; there is no such hierarchy.[6] However, it is not possible for Velleman to endorse a non-hierarchical relation between love and respect while preserving his answer to the second objection on selectivity. This is because, as seen above, an empirical persona—the immediate object of love's evaluation—is an imperfect representation of a person's rational nature—the final object of evaluation. Harcourt and Bagley are correct: if empirical personae are imperfect representations, it does follow that, if we were able to offer perfect representations of our rational nature, then our observable qualities would be superfluous in love's evaluation. Hence, Velleman's answer to selectivity and to the personal character of love are incompatible with each other.

Due to this incompatibility, Velleman does not succeed in putting forward an account of love that is personal and selective. For Harcourt and Bagley, their objections mean that Velleman's claim that rational nature is the final object of evaluation is incorrect: this is the main flaw they see in the account. Harcourt (2009, 352) formulates this conclusion explicitly: "Love is not to be analysed as a response to the very same value as also elicits respect, namely, the value we have in virtue of exemplifying rational nature." However, they miss a second possibility available for Velleman which does not require giving up rational nature as the final object of evaluation. A non-hierarchical relation between love and respect could be preserved by re-defining the immediate object of evaluation. To avoid the hierarchical relation that leads to impersonal love, we need to re-interpret empirical personae so that they are not imperfect representations of rational natures. That would preserve Velleman's solution for selectivity while avoiding the incompatibility between Velleman's responses to the personal and selective characters of love.[7]

I will show next that this alternative becomes available by substituting the Kantian framework in the current formulation of the view with narrative theory—in fact, Velleman's own narrative theory. Specifically, I argue that if we replace 'rational nature' with 'inner narrator' as love's final object of evaluation, and 'empirical persona' with 'observable narrative' as love's immediate object of evaluation, the resulting re-formulated view can overcome the objections I have presented in this section.

6 At one point Velleman (1999, 366) hints at a hierarchy that goes in the opposite direction: "I regard respect and love as the required minimum and the optional maximum responses to one and the same value." This can also be understood to be implicit in his differentiation between "*merely* merely respecting rather than loving" (372; my emphasis). I do not think this claim is seriously supported in the paper, so I ignore it here.

7 A third possibility would be to bite the bullet for the second objection. Kieran Setiya (2014) adopts this strategy in order to put forward a view that mirrors Velleman's answer to the partiality challenge, but which abandons the Kantian framework. For Setiya, the final object of evaluation in love is the property of humanity, and the immediate object of evaluation is contingent—that is, he embraces Bagley's criticism of selectivity being random.

3. Velleman on Love: A Re-Formulation

In its current formulation, Velleman's answer to the descriptive question about love seems inseparable from (his interpretation of) Kantian moral theory, given that rational natures are the final object of evaluation. But it is possible to isolate the view from its Kantian element and frame it in a normatively neutral way. Velleman says that love is a mode of valuing another person that results in the arresting of the lover's emotional defences. The value of the loved person is apprehended through her observable qualities—ranging from personality to physical traits and mannerisms—and ultimately directed at that person's true self, which is represented in, and perceived through, the loved person's observable qualities. The self is not to be understood as a distinguishable entity: the true self *is* the capacity to act for reasons—rational nature (see footnote 2). Given that I am presenting Velleman's view minus the Kantian framework, let us substitute "rational nature" for a neutral term—agency—which is widely understood as the capacity to act for reasons. To sum up, love consists in an evaluation of the observable qualities of a person *qua* representations of that person's agency, which results in the lowering of the lover's emotional defences. Loving a person is, then, loving an agent.

Velleman's work on agency and practical rationality spans across four decades. Capturing it in its entirety would be a fruitless enterprise—and a rather unnecessary one. What matters for my purposes is that Velleman formulated the "capacity to act for reasons" or "the self" in terms that—as I will argue later—allow for an alternative route to understanding love as an evaluation of agency. Here I focus specifically on Velleman's explanation of how the different formulations of the self he has put forward across his different works interact and relate with each other. I focus on Velleman's reply to Catriona Mackenzie (2007), who was the first to highlight the interrelatedness between different aspects of the self in her review of *Self to Self* (Velleman 2006).[8]

Velleman discusses the self from three different perspectives: as an autonomous agent, as a diachronically unified entity, and an inner narrator. As Mackenzie (2007, 263–64) points out, these are not oppositional or distinct views of the self, but descriptions of "different dimensions or aspects of selfhood." These three dimensions of the self are all modes of self-reflection; they capture "interactions between the person and himself under various guises" (Velleman 2007, 284). I leave aside the dimension of diachronic unity, which is tangential to my re-formulation of Velleman's view on love.[9] With that in mind, I proceed on the understanding that the autonomous agent and the inner narrator are dimensions of the self that constitute two different modes of self-reflection.

Velleman's view of love rests on the first dimension of the self as an autonomous agent, understood in the Kantian sense as an instance of rational nature. Given that both the autonomous agent and the inner narrator are dimensions of a person's self, and that, for Velleman, love is an evaluation of a person's self, would it be possible to replace "rational nature" with "inner narrator"?

8 Velleman (2007, 284) agrees with Mackenzie that he had previously overlooked the relation between the different dimensions of the self. I am truly indebted to Mackenzie's discussion of Velleman's views, which, together with Velleman's responses to her, often offer much clearer insights than the original works.

9 This does not mean that diachronicity is irrelevant to love as a phenomenon (cf. Jones 2008, Schechtman 2021, or Jollimore 2022), or that Velleman's view on diachronic unity may not be useful in understanding important aspects of love. However, my purpose here is to overcome the objections set up in Section 2, which are not connected to diachronicity.

Velleman refers to a person's rational nature as her capacity to act for reasons, and he refers to the inner narrator in the same terms (2005, 69). The inner narrator is a person's capacity to interpret events in her life (her behaviours, interactions, thoughts, and occurrences around them) in the form of narratives, which in turn give her reasons to act in certain ways and not others. Narratives are not to be understood as fully-fledged stories, but as the structures that our behaviours and mental states take in following from one another, insofar as they are guided by our inner narrator. For example, if I decide I am going to have some avocado on toast for dinner, I have already set an aspect of how the story of me is going to develop today (it will involve, at some point, having avocado on toast for dinner, but also taking avocado out of the fridge, setting out the appropriate cutlery, or making sure the toaster's settings are not too high, for example). This is not because by deciding to have avocado I have somehow determined what my future actions will be. Instead, I have set out the conditions of intelligibility for my future actions in light of those plans. I am still free to choose to take the bacon out of the fridge instead of the avocado, but that action would be incoherent with my intention of dining on avocado on toast. According to Velleman, the reason why I choose the action that will cohere with my existent intentions and mental states is because the aim of action itself is intelligibility. This is because when we act for reasons we are aiming to make sense of ourselves: intelligibility is the constitutive aim of action.[10] Whether a person has a reason for taking an action "ultimately depends on whether it would make for a coherent continuation of his story" (69). The autonomous agent and the inner narrator have the same role (acting for reasons) and the same aim (intelligibility), and differ only insofar as they constitute different modes of engaging with oneself, as seen above. On that basis, a person's rational nature is plausibly interchangeable with her inner narrator as the final object of evaluation of love, since they are analogous aspects of agency. As I argue next, this exchange allows Velleman to overcome the objections set up in Section 2, with a re-formulated view of love that is both personal and selective.

I said at the end of Section 2 that the main flaw of Velleman's view—which opens it up to the objections seen above—is his understanding of the immediate object of evaluation as an imperfect representation of a person's true self, and not the fact that a person's rational nature is the final object of evaluation. Now I can make that claim more precise: the main flaw in Velleman's view on love is that rational nature is the wrong dimension of agency to focus on as the final object of evaluation, and that this mistake leads him to an incorrect understanding of the immediate object of evaluation as an imperfect representation of a person's true self. Let us see how focusing on the inner narrator as the final object of evaluation instead of rational nature helps Velleman overcome the objections from Section 2.

Again, the first objection is that the hierarchical relation between love and respect that Harcourt and Bagley identify in Velleman's view means that love collapses into respect, and thus Velleman does not account for the personal character of love. This criticism is inescapable insofar as Velleman says that the empirical persona—the observable qualities that are the immediate object of evaluation—is an imperfect representation of a person's agency. Replacing rational nature with inner narrator opens up a new way of understanding the empirical persona—not as an imperfect representation, but as a *product* of agency. I unpack this next.

If a person's inner narrator is her capacity to act for reasons, then *actions* are the product of the work of the inner narrator. These actions are themselves the result of intentions, desires, and beliefs that the inner narrator interprets as intelligible, that is, understands as belonging to itself: these *interpretations* are also the

10 Velleman makes different formulations of the constitutive aim of action. Here I am drawing from his "Précis of the Possibility of Practical Reason" (2004) where he explains how his view evolved from the Kantian view of practical reason that he had defended thus far (including in the eponymous paper included in that collection).

product of the work of the inner narrator. When we apprehend a person's traits, we are perceiving that person's actions and interpretations in interaction—which in turn are expressions of those traits. What that person does, and how that person understands the world (if they are impulsive, kind, or vengeful), are the direct product of her aiming at intelligibility through her inner narrator. Let us call this collection of a person's actions and interpretations her "observable narrative." An observable narrative is plausibly constituted by the same elements as the empirical persona, as initially formulated by Velleman: the manifest person, which is the collection of qualities of the loved person that are perceived by the lover, and the motivation for the arresting of the lover's emotional defences.

Because an observable narrative is a product of the inner narrator, and not a representation of the inner narrator, the obstacle to Velleman's aim of incorporating the personal character of love goes away.[11] Recall that Velleman states that love and respect are two different modes of evaluation which stand in a non-hierarchical relation with each other. Velleman's claim that an empirical persona is an imperfect representation of the final object of evaluation allowed Harcourt and Bagley to argue that love is a "poor" interpretation of the incomparable value of a person, or a "mistake" that should be corrected. But if the immediate evaluation is not a poor interpretation or a mistake, then we can accept Velleman's claim that love is not a lower kind of evaluation than respect. If the final object is the inner narrator, the immediate object—the observable narrative—is no longer a distorted reflection of the agent's value. It is actually the product of her agential powers. Hence, Velleman's account can be an account of personal love—of love for particular individuals—despite the final object of evaluation being something that everyone has—agency.

This re-formulated view does not yet avoid a hierarchical relationship between love and respect, as becomes evident when attempting to answer the question of selectivity. Recall that saying that love is selective means that love for a person is based on something that sets her apart from others, in such a way that that person is loved because of something in them that distinguishes them from others. For Velleman, we love some people and not others because we have different skills in performing our own value through our empirical personae and perceiving other people's value through their empirical personae. Again, if we isolate the view from its Kantian framework, this is not an implausible claim: when we say we want to be loved because we are special, what we actually mean is that we want to be loved because we are *seen* as special, or because we are special *for the lover*. In that sense, differences between people as objects of love are constituted not only by the characteristics of a specific individual as an object of love, but by the perception of those characteristics by the lover. The problem in Velleman's original proposal is that this difference in perception is formulated as a defect—we are imperfect performers and perceivers of value. As long as this is the case, it is not possible to overcome the first objection. So the answer to selectivity needs to account for differences in perception of value without this difference being defective in nature. Again, an alternative answer to the question of selectivity becomes available by applying Velleman's (2003) own insights on narrative explanation.

According to Velleman, when putting together events in narrative form we are not just giving a causal story of what has happened, is happening, or is going to happen. Instead, it is the emotional cadence or sense of anticipation for what comes next that turns a series of events into a narrative (6). That is, the sequence of events needs to prompt an emotional response in the audience. This emotional cadence is sustained throughout the duration of the narrative—although the actual individuated emotions an audience feels when engaging with a narrative may change during the plot, and the audience might go from fear to anticipation to relief, and so

11 This may be disputed by the claim that people may "see through" observable narratives or "buy into" inaccurate observable narratives. I respond to this objection to the re-formulated view in Section 4.

on. As Andrew Speight points out, “For Velleman, narrative helps us assimilate events not to a pattern of how things happen but to a pattern of how things feel or what they mean—an emotional cadence that, on his view, connects with the biological organism’s desire for tension and release” (2014, 54). This emotional reaction and its power to give meaning to isolated events answer for the “distinctive force that narratives have on us” (53). Different people find different narratives meaningful, and this influences all types of choices that people make on the basis of, and as an expression of, the things that they care about. For example, someone may pick a career in medicine because a “saving lives” narrative is meaningful to them, while others may dedicate their time and effort to winning an Olympic medal or climbing the highest peaks in the world because of an “overcoming challenges” narrative.

However, narratives need not be explicit to elicit emotion. When the inner narrator aims to make the world intelligible to itself, it is often confronted with different possibilities and events that are more intelligible than others. To achieve intelligibility is not merely to form a propositional belief that something is intelligible. Things being intelligible *feels* like something: it feels like things make sense. In previous work (Lopez-Cantero 2023) I have drawn from Matthew Ratcliffe’s (2016) view on existential feelings to describe the feeling of intelligibility as a shift in one’s experience of the world, enabled by the opening of a new branch of possibilities that opens when one falls in love. It is this experience of intelligibility or meaningfulness, which for Velleman is an experience of value, that allows for the arresting of the lover’s emotional defences. The existential feeling of intelligibility becomes a background orientation once one is in love: for as long as one is in love with a particular individual, loving that individual is one aspect that makes the lover’s life meaningful. That does not mean that loving that particular individual always makes the lover’s life happier, or more filled with meaning, understood as self-realisation: as Velleman (1999, 361) says, “love also lays us open to feeling hurt, anger, resentment, and even hate.” It means that, without that particular individual, the lover’s life would be less intelligible to themselves as theirs.

The answer to the question of selectivity, then, lies in the fact that different people find different narratives meaningful, which means that they will experience the existential feeling of intelligibility when confronted with different possibilities—that is, different people. We love some people and not others because some observable narratives make more sense to us than others—they feel more meaningful or more intelligible. That is the explanation for the selective character of love. My response to the question of selectivity has a double payoff. On one hand, it allows for a re-formulation of Velleman’s view that wholly avoids the objections set out in Section 2, while preserving the phenomenological component of love. On the other, the re-formulation makes apparent that love engages the lover’s agency as well as the loved person’s agency, as I explain next.

It could be objected at this point that my proposal suffers from the same flaw as the original one: one may say that if we were able to perceive the meaningfulness of everyone’s observable narratives, then we should rationally love everyone—that is, that love is a “mistake” on the side of the lover. However, such an objection would result from an oversight of the constitutive role of the lover’s inner narrator in love’s evaluation. The fact that a person finds some narratives more meaningful than others is not an inevitable consequence of our limited skills at perceiving meaning: it is, in fact, an expression of those skills. What we find salient in the first place has to do with our way of understanding the world, and insofar as have a specific way of understanding the world, our inner narrators are successful when they succeed in guiding actions and attitudes in a way that is intelligible to ourselves. This means that the inner narrator needs to discriminate between what is perceived and interpreted by default: it is the capacity to do precisely that. For example, a person who is an avid gym-goer may notice and appreciate people’s level of physical fitness in a way that a bookworm academic may be

completely oblivious to. This difference in perceiving the world will also influence how each of those individuals understands a first encounter with other individuals in virtue of their levels of fitness. And that, in turn, will influence their relationships with that person, as well as the extent to which that person elicits the feeling of intelligibility. For those reasons, selectivity in love is then not a matter of poor performance of one's skills, as Velleman's initial formulation states. Instead, it is a natural result of the functioning of the lover's inner narrator.

From the above, we now have a fully-fledged new view of love as an evaluation of a person's agency which avoids the objections set out in Section 2. We meet a person with yellow hair we consider beautiful, a mind we consider brilliant, and a sense of humour that we find amusing. We do not consider those things separately, but we give them unity as part her observable narrative, which is the immediate object of loving evaluation through which we perceive the final object of loving evaluation, the inner narrator. Love is an evaluation of something that is unique to the beloved, and not of something that everyone has. What is more, it is not an imperfect representation of her value as an agent, but a direct product of her agency. In that way, the objections based on the personal and selective characters of love are overcome. Further, the re-formulated view gives prominence to the lover's agency as well. As agents, we aim at intelligibility through our inner narrators, and whether we experience meaningfulness when confronted with a person's observable narrative will depend on our own way of understanding the world.[12]

In this way, I have offered a way to re-formulate Velleman's agential view, drawing from his own claims about narrative, while at the same time addressing the objection that it is an impersonal account, and renouncing the Kantian framework that made the account counterintuitive to start with. The re-formulated view, however, may be exposed to new objections. I finish the paper by briefly addressing three of these objections.

4. Objections: Narrativity and Morality

It could be argued that, by giving up Velleman's Kantian framework, the re-formulated view runs into three new problems in terms of its commitment to narrative theory, and in terms of the compatibility of love with moral action.

It goes without saying that the re-formulated view would be unpalatable to narrative sceptics who reject the existence of an inner narrator wholesale. However, such levels of scepticism are rare. Reasonable anti-narrativism is often formulated in more tempered ways, in the sense that the inner narrator is not a universal capacity that everyone has, or in the sense that it is not a capacity that we should cultivate. Non-generalisation seems like a potential problem for the view of love I have put forward: if not everyone has the capacity for inner narration, can people without an inner narrator love at all? Galen Strawson, for example, has repeatedly proclaimed himself an "episodic," insofar as he claims to not have an inner narrator. In fact, Strawson (2007, 109) explicitly addresses how love may happen without the intervention of the inner narrator:

12 In a previous version of this manuscript, I used the expression *narrative fit* as part of my answer to the question of selectivity. This is the exact quote: "Given that the lover perceives other people's personal narratives in relation to her own personal narrative, some of these narratives will be more meaningful, or have a higher promise of meaning, than others. There will be what I call a narrative fit." Lotte Spreeuwenberg (2021) criticises this view as giving too passive a role to the loved person. I can see how the expression and the wording of the original manuscript could lead to this misunderstanding, and I have left the notion of a "narrative fit" aside, to be explored in further work as an explanation of mutuality instead of selectivity. I am thankful to the reviewer who made me realise the potential confusion between selectivity and mutuality that stems from the notion of narrative fit. Given my argument in this section about the role of both the loved person's and the lover's agency, I think my proposal avoids Spreeuwenberg's criticisms.

> Enduring love of a person is, at any moment, a matter of present disposition. Its manners and customs may be shaped by the past, but it does not require any tendency to engage in explicit recollection of the past, nor any trace of any Diachronic sense that one—or the one one loves—was there in the past. ... A gift for friendship doesn't require any ability to recall past shared experiences in any detail, nor any tendency to value them. It is shown in how one is in the present.

Strawson's statement reflects a mistake frequently made by anti-narrativists, which is to conflate intelligibility with diachronic understanding. In the previous section, I explained that Velleman distinguishes three dimensions of the self: agent, inner narrator, and diachronically unified entity. Strawson's objection is directed at the third dimension of the self, which I said I would not discuss here, since the objections based on the personal and selective character of love are not directly related to diachronicity. The inner narrator aims at intelligibility and not diachronic unification. Strawson does not deny that agents aim at intelligibility; he merely denies that the capacity to make one's life intelligible is that of an inner narrator. But his dispute here is terminological: he believes that the capacity of intelligibility is only trivially narrative (2007 440). This paper is not, however, the place to solve the long-standing dispute between narrativists and anti-narrativists. I am happy to accept that anti-narrativists may not fully accept my view, but, as their view stands, it is not a rejection of the universal character of the capacity of intelligibility that is at the heart of the re-formulated view. I think this is enough to move on from this potential objection.

The potential dangers and limitations of the inner narrator, on the other hand, do seem like direct problems for my account. Because the inner narrator interprets the world from its own internal perspective, it is possible that these interpretations are unreliable. This is a reasonable worry for my account of love. We have all had a friend whose romantic partner is the most obnoxious human being we have ever encountered, and have sat though endless eulogies from our friend about how attentive, kind, and funny their partner is. In those cases, we may be the ones in the wrong if we fail to see something that is really there, but we could also have evidence that goes to show that our friend is the one has "bought into" their partner's observable narrative. Sometimes, the misinterpretation goes in the other direction: the famous trope of enemies to lovers illustrates how initial interpretations of the other may be inaccurate, and that the person who seemed unbearable to begin with reveals herself as an adorable grump. Further, sometimes we are unreliable as inner narrators of our own lives. We may present an observable narrative that portrays us as arrogant and nonchalant when in reality we are sensitive and insecure, and the people who love us are the best positioned to "see through" our observable narratives. These seem to be problems that arise directly from the incorporation of narrative theory, and which were not present in Velleman's initial proposal. Velleman already accounted for unreliability by arguing that we are poor performers and interpreters of value, but since I have rejected that the imperfect character of the immediate object of love's evaluation, this answer is not available in the re-formulated view.

Again, recall that the observable narrative is a product, and not a representation, of the inner narrator. The inner narrator is just a capacity, not a real self that needs to, or can be, uncovered through engagement with the observable narrative. The notion of unreliability should not be understood as lack of skill (intentional or not) at one's portrayal or apprehension of the final object of evaluation. Instead, it is better to understand the scenarios above as instances where the lover accesses alternative interpretations of the loved person's observable narrative through her own inner narrator. Dean Cocking and Jeanette Kennett (1998) argue, in fact, that interpreting the actions and attitudes of the people we love is a constitutive and necessary element of love; in turn, the loved person must be open and receptive to be shaped by the lover's interpretation. Of course, interpretation can be taken too far, as in the case of the friend with the obnoxious partner. Philosophers of love have often worried about this so-called problem of imagined qualities: all cognitivist views that give a role to

the characteristics of the loved person potentially suffer from this problem. In that sense, even if it is a potential objection to the view I have presented, it is an objection insofar as the view gives a role to the loved person's characteristics. The only alternative to avoid the objection would be, then, to say that the loved person's characteristics do not matter, which as we have seen results in the problem of the potential interchangeability of the loved person. But also, the problem of imagined qualities has been greatly exaggerated. Here I side with Troy Jollimore (2011, 7): even if the lover's apprehension of the loved person's qualities is not completely accurate, in most cases there will be sufficient accuracy, and cases where all characteristics are imagined and yet there is actual love are rather unusual. In that sense, the limitations of the inner narrator are not worrying enough to abandon the re-formulated view.

A second objection is that, by re-formulating the view in terms of narrative theory, it becomes apparent that the structure of love having an immediate and a final object of evaluation is not needed. Indeed, I have shown that the observable narrative is the product of the inner narrator, but I have merely assumed that the inner narrator is the final object of love's evaluation. Maybe it would be better, and more conceptually economical, to get rid of the final object altogether, and see the inner narrator simply as the cause or the driver for love, but not as the object of love's evaluation. I think that could be a plausible conclusion from my argument. Still, arguing that love is an evaluation of the loved person's observable narrative that arrests the lover's emotional defences, and that engages both the lover and the loved person's agency, seems like a worthwhile aim to me, so not all would be lost even if I bit that bullet. Whether the bullet needs to be bit or not depends on the answer to a third and final objection that I explain next.

At first glance, the re-formulated view seems to completely abandon Velleman's primary aim, which was to show that love is a moral emotion. It could even be argued that Velleman's descriptive aim (to define what love is) is subordinate to his normative aim (to show that love is not at odds with moral action). In that sense, it seems plausible that Velleman, or someone who endorses his original account, would reject the re-formulated view, which has been stripped of the Kantian element that allowed the compatibility between love and moral action. Recall that Velleman argues that love can be incorporated into moral consideration because love does not discriminate between people in terms of value, given that it is an appreciation of the loved person's rational nature.

On a second look, however, it becomes clear that the re-formulated view is wholly compatible not only with Velleman's interpretation of Kantian ethics, but also with other normative approaches. Starting with Velleman, the re-formulated view is still compatible with the claim that love does not discriminate between people in terms of value. Rather, it discriminates between people in terms of meaningfulness: the fact that some observable narratives elicit the feeling of intelligibility in us has no bearing on whether the people with those observable narratives are more or less valuable. Importantly, the fact that the re-formulated view is morally neutral but agency-based means that it is possible to fill in the details on the moral character of love with normative theories other than (Velleman's) Kantian theory. As I claimed in Section 1, the re-formulated view is a valuable contribution to the wider debate on love as an agential process. Unless we want to separate agency from moral action, the structure of the immediate and final object of evaluation has to be preserved: agency should be constitutive of love, and not merely its driver. Hence, the answer to the second objection—whether the final object of evaluation is superfluous—depends on whether we want the account to be compatible with a normative claim about love—that love is compatible with moral deliberation. If this compatibility is to be preserved, it is not possible to forego the structure of love's evaluation in the re-formulated view. Those who do not need to preserve this compatibility—because they do not deem it necessary for a descriptive view or

because they do not believe that love and morality are compatible—may be able to do away with agency as the final object of evaluation. I do not take a side here: the view is formulated so it can be adapted to a variety of normative projects, instead of being tied to a specific understanding of Kant, which plausibly was the root of the widespread resistance to Velleman's original proposal all along.

5. Conclusion

The re-formulation of Velleman's account I have offered here can open new avenues to discussing love as an agential process, where agency is not only in loving but in being loved. This is not to be understood in a highly intellectualised way, as if agents were only in the business of applying reasons to actions and deciding on the morally optimum choice. Instead, agents—people—aim to understand the world and themselves. In the re-formulated account, this capacity to make the world intelligible is put at the centre of loving and being loved. Love's final object of evaluation is the inner narrator, while the immediate object is the observable narrative, itself perceived—and shaped—by the lover, through the work of their own inner narrator. Velleman (1999, 365) says that, in love, "what our hearts respond is to another heart"; the re-formulated view preserves that claim, and it reveals that when loving a particular other we are making sense of the world with others—or at least trying to.

Acknowledgements
I am grateful to the organisers and participants of the Inter-University Workshop on Philosophy and Cognitive Science that took place at Carlos III University (Madrid) in 2017, where I first presented my idea. I am also very grateful to the people who gave useful and encouraging feedback to earlier drafts—particularly David Velleman, Edward Harcourt, Helen Beebee, Thomas Smith, and Lotte Spreeuwenberg—as well as two anonymous reviewers, whose comments significantly contributed to the improvement of the final version.

References

Abramson, K. and A. Leite. 2011. "Love as a Reactive Emotion." *The Philosophical Quarterly*, 61 (245): 673–99.

Bagley, B. 2015. "Loving Someone in Particular." *Ethics* 125 (2): 477–507.

Candiotto, L. and H. De Jaegher. 2021. "Love In-Between." *The Journal of Ethics* 25 (4): 1–24.

Chappell, S. G. 2014. *Knowing What to Do: Imagination, Virtue, and Platonism in Ethics*. Oxford University Press.

Cocking, D. and J. Kennett. 1998. "Friendship and the Self." *Ethics* 108 (3): 502–27.

Ebels-Duggan, K. 2008. "Against Beneficence: A Normative Account of Love." *Ethics* 119 (1): 142–70.

———. 2019. "Love and Agency." In *Routledge Handbook of Love in Philosophy*, edited by A. Martin, 300–12. Routledge.

Foster, G. 2018. "Persons and Properties: A Sartrean Perspective on Love's Object." *European Journal of Philosophy* 26 (1): 82–94.

Frankfurt, H. 2004.*The Reasons of Love*. Princeton University Press.

Han, Y. 2021. "Do We Love for Reasons?" *Philosophy and Phenomenological Research* 102: 107–27.

Harcourt, E. 2009. "Velleman on Love and Ideals of Rational Humanity." *The Philosophical Quarterly* 59 (235): 349–56.

Helm, B. 2010. *Love, Friendship, and the Self: Intimacy, Identification, and the Social Nature of Persons*. Oxford University Press.

Jeske, D. 2017. "The Rational Significance of Intimate Relationships." In *Love, Reason and Morality*, edited by E. Kroeker and K. Schaubroeck, 75–92. Routledge.

Jollimore, T. 2011. *Love's Vision*. Princeton University Press.

———. 2022. "Love in the Past (and Present, and Future)." In *Philosophy of Love in the Past, Present and Future*, edited by A.

Grahle, J. Saunders, and N. McKeever, 89–106. Routledge.
Jones, K. 2008. "How to Change the Past." In *Practical Identity and Narrative Agency*, edited by K. Atkins and C. Mackenzie, 269–288. Routledge.
Keller, S. 2000. "How Do I Love Thee? Let Me Count the Properties." *American Philosophical Quarterly* 37 (2): 163–73.
Kennett, J. 2008. "True and Proper Selves: Velleman on Love." *Ethics* 118 (2): 213–27.
Kolodny, N. 2003. "Love as Valuing a Relationship." *Philosophical Review* 112 (2): 135–89.
Kroeker, E. 2019. "Reasons for Love." In *Routledge Handbook of Love in Philosophy*, edited by Adrienne Martin, 277–287. Routledge.
Kühler, M. 2014. "Loving Persons: Activity and Passivity in Romantic Love." In *Love and Its Objects: What Can We Care For?* edited by C. Maurer, T. Milligan, and K. Pakovská, 41–55. Palgrave Macmillan.
Lopez-Cantero, P. 2023. "Three Stages of Love, Narrative, and Self-Understanding." In *Emotional Self-Knowledge*, edited by A. Montes Sánchez and A. Salice, 147–67. Routledge.
Mackenzie, C. 2007. "Bare Personhood? Velleman on Selfhood." *Philosophical Explorations* 10 (3): 263–82.
Naar, H. 2017. "Subject-Relative Reasons for Love." *Ratio* 30 (2): 197–214.
Ratcliffe, M. 2016. "Existential Feeling and Narrative." In *Funktionen des Lebendigen*, edited by T. Breyer and O. Müller, 169–92. Walter de Gruyter.
Saunders, J. and N. McKeever. 2022. "Irrational Love: Taking Romeo and Juliet Seriously." *International Journal of Philosophical Studies* 3: 254–75.
Schaubroeck, L. 2014. "Loving the Lovable." In *Love and Its Objects: What Can We Care For?* Edited by C. Maurer, T. Milligan, and K. Pakovská, 108–124. Palgrave Macmillan.
Schechtman, M. 2021. Time and Transcendence in the Before trilogy. In *Before Sunrise*, *Before Sunset*, *Before Midnight: A Philosophical Exploration*, edited by H. Maes and K. Schaubroeck, 24–40. Routledge.
Setiya, K. 2014. "Love and the Value of a Life." *Philosophical Review* 123 (3): 251–80.
Speight, A. 2014. "The Narrative Shape of Agency: Three Contemporary Philosophical Perspectives." In *Narrative, Philosophy and Life*, edited by A. Speight, 49–60. Springer.
Spreeuwenberg, L. 2021. "'Love' as a Practice: Looking at Real People." In *New Philosophical Essays in Love and Loving*, edited by S. Cushing, 63–86. Palgrave Macmillan.
Strawson, G. 2007. "Episodic Ethics." *Royal Institute of Philosophy Supplement* 60: 85–116.
Velleman, J. D. 1999. "Love as a Moral Emotion." *Ethics*, 109 (2): 338–74
———. 2003. "Narrative Explanation." *The Philosophical Review*, 112 (1): 1–25.
———. 2004. "Précis of the Possibility of Practical Reason." *Philosophical Studies* 121: 225–38.
———. 2005. "The Self as Narrator." In *Autonomy and the Challenges to Liberalism: New Essays*, edited by J. Christman and J. Anderson, 56–76. Cambridge University Press.
———. 2006. *Self to Self: Selected Essays.* Cambridge University Press.
———. 2007. "Reply to Catriona Mackenzie." *Philosophical Explorations* 10 (3): 283–90.
Yeats, W. B. 1956. *The Collected Poems of W. B. Yeats*. Macmillan.
Zangwill, N. 2013. "Love: Gloriously Amoral and Arational." *Philosophical Explorations* 16 (3): 298–314.

This work is licensed under a Creative Commons Attribution 4.0 International License.