BEING IN LOVE:
A NARRATIVE ACCOUNT

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2019

Pilar Lopez Cantero

School of Social Sciences
Department of Philosophy
Table of contents

Table of contents ........................................................................................................ 3
Abstract ...................................................................................................................... 5
Declaration .................................................................................................................. 6
Copyright statement .................................................................................................. 7
Acknowledgements ..................................................................................................... 8
Dedication ................................................................................................................... 11
Introduction ................................................................................................................ 12

Chapter 1. Love, caring, and identity ................................................................. 16
  1.1. The basics of caring .................................................................................... 17
      1.1.1. Frankfurt on love and identification ............................................... 17
      1.1.2. Helm’s critique ............................................................................ 19
  1.2. Helm’s intimate identification .................................................................. 20
      1.2.1. Caring about persons .................................................................... 20
      1.2.2. Identity, well-being and a life worth living .................................. 24
      1.2.3. Problematic identities .................................................................. 27
  1.3. Rejecting the caring approach .................................................................. 30
      1.3.1. Love in spite of problematic identities ........................................ 30
      1.3.2. Love in spite of opposing identities ............................................. 34
      1.3.3. Not caring about neutral values .................................................... 36
  1.4. Making love about the lover ....................................................................... 40
Conclusion .................................................................................................................. 42

Chapter 2. Narrative theory ............................................................................... 43
  2.1. Preliminary distinctions .............................................................................. 44
      2.1.1. Departing from the philosophy of emotion .................................... 44
      2.1.2. Departing from the metaphysics of personal identity .................... 46
      2.1.3. The self-concept .......................................................................... 48
  2.2. Different formulations of the narrative claim ............................................ 52
      2.2.1. The basics of narrative .................................................................. 52
      2.2.2. Strong Narrativism ...................................................................... 54
      2.2.3. Minimal Narrativism .................................................................... 57
  2.3. The anti-narrativist challenge ..................................................................... 61
      2.3.1. Accepting non-diachronic self-experience .................................... 62
      2.3.2. Accepting non-narrative self-understanding ................................ 64
      2.3.3. Qualifying story-telling .................................................................. 65
      2.3.4. Leaving revision for later ................................................................ 68
  2.4. The challenge to Minimal Narrativism ....................................................... 69
Conclusion .................................................................................................................. 72

Chapter 3. The narrative self-concept ............................................................... 73
  3.1. The role of narrative in Minimal Narrativism ............................................ 74
      3.1.1. Maximal understanding ................................................................. 74
      3.1.2. Narrative explanation vs causal explanation .................................. 76
      3.1.3. Hutto’s narrative models ................................................................. 79
      3.1.4. Goldie’s meaningfulness and import .............................................. 81
Abstract

This thesis defines a feature of romantic love: what being in love is. My claim is that being in love is a defining trait of the narrative self-concept, which articulates a person’s understanding of herself and the world.

In chapter 1, I start from a plausible answer to what being in love is: a specific kind of caring about someone. I show that the caring view, exemplified by Frankfurt (1998, 2004) and Helm (2010a), does not accommodate cases of non-harmonic love. Hence, I suggest a change of direction. Instead of focusing on the attitude that the lover has towards the loved person – caring – I propose to focus on the fact that being in love entails a change for the lover. Being in love entails a change in the lover’s identity: that is the starting point for my view.

In chapter 2, I situate the view at the level of the self-concept and put forward narrative theory as the appropriate framework for my purposes. I distinguish between two formulations of narrative theory: Strong and Minimal Narrativism. I present the anti-narrativist challenge set by Strawson (2004) and argue that a plausible narrative theory of the self-concept can only be formulated within Minimal Narrativism.

In chapter 3, I give my account of the narrative self-concept. From Hutto’s (2016) and Goldie’s (2012) views, I establish that narrative is a route for ‘maximal understanding’ of reasons for action. I show that there is a missing piece in Minimal Narrativism: salience. Salience solves both problems and grounds the feedback loop between self-concept and action: people’s beliefs about themselves shape what they do, which in turn influences these beliefs. This feedback loop is diachronic, given that current self-concepts are partly constituted by past self-concepts.

In chapter 4, I draw from Cocking and Kennet (1998), Rorty (2016) and Nehamas (2010) to define the ‘mutual shaping view’. According to the mutual shaping view, love requires openness to have one’s self-concept shaped by the loved person. This process is what Jones (2008) calls trajectory-dependent, i.e., a narrative. In the mutual shaping and the trajectory views, however, love seems to justify itself. This problem is solved by introducing self-concept consistency as a defeasible reason.

In chapter 5, I offer my account: the ‘trait view’ of being in love. This approach accommodates the best intuitions of the views I discuss throughout the thesis: it explains the change in the lover’s self-concept while giving the loved person an important role in the lover’s conception of a life worth living. I argue that what ‘being in love’ is will vary between individuals, and thus is agent-dependent. Finally, I briefly deal with the objection of whether this account explains what being in love is for self-proclaimed non-narratives, who do not have narrative self-understanding. Non-narratives, I argue, can be in love; but they can never acquire maximal understanding as to what is, for them, to be in love.
Declaration

I, Pilar Lopez Cantero, hereby declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
Acknowledgements

Over the past four years, my self-introduction to strangers with “I am doing a PhD in Philosophy” has often been met with a part-bemused, part-pitiful remark of “I wouldn’t be able to do that”. Then I would follow with an “It’s on philosophy of love” which would provide with an ice-breaker and prevent me from going into an emotional diatribe. The truth is that what has made me able to write this thesis is having an incredible group of people in my professional and personal life. I have been saving the emotional diatribe for this Acknowledgements section.

First and foremost, I thank my supervisors, Tom Smith and Joel Smith –whom my dad nicknamed ‘the Smiths’ straight away, despite probably not knowing about the music band. Tom and Joel are the perfect team: their different but complementary supervisory styles have made me the best philosopher that I could be at this stage. Thanks, Joel, for pushing me to learn the rules of the game and for dealing with my subsequent million questions about academia. And thanks, Tom, for having seen the exact pace I needed, and asking for progress without putting unnecessary pressure on me, while always taking on board my ideas for long and interesting discussions. Your supervision style should be in any PhD mentor guidelines.

The thesis has benefited immensely from the comments of Helen Beebee, Katrien Schauwbroeck, Hans Maes, Edward Harcourt, David Velleman, Maureen Sie, and Marya Schechtman; as well as the contributions from participants at the meetings of the Society of Applied Philosophy and the University of Antwerp’s Summer School The Right to Be Loved –specially Michäel Van Remoortere’s insights on life, narrative and Taylor Swift.

I am also grateful to the up-and-coming group of young philosophers of love formed by Lotte Spreeuwenberg, Natasha McKeever, Luke Brunning, Alison Toop and Joe Saunders. Like Lotte says, “Love is hard work!!!”, but it’s made easier by having this network of great people sharing that work with you. Thanks, Lotte, for providing with that quote, for our gossip dinners and for our discussions about how our theories, when merged, will solve Philosophy.

Also, I have been lucky to count with an all-star team of proof-readers: Lydia Farina, Clint Verdonshot, Janet Hanley, Dave Maison, Simon Walgenbach and Danny Coposescu who all used their personal time to fix mistakes and leave hilarious comments on the margins, making editing much more entertaining.

And, of course, let us not forget about money: I wouldn’t have been able to afford to do this amazing project without the funding of the School of Social Sciences, the Society of Applied Philosophy and the Aristotelian Society.

The support I have received as an academic goes all the way back to my master’s time in Edinburgh. My master’s supervisor, David Levy, showed me that I didn’t need to change my way of thinking into ‘factory thinking’ and enthusiastically supported my project. Also, Alfred Archer has gone from Edinburgh colleague to
friend, co-author and academic agony aunt. More importantly, having entered the field with precarious knowledge and a crippling fear of terms like ‘internalism about reasons’ or ‘supervenience’, I would not have made it to the PhD without the absolutely fantastic group of people that I was lucky to meet there: Giorgos Seretis, Guillermo Valverde, Franklin Jacoby, Becky Millar, and specially Danny Coposescu and Melanie Sarzano. Danny took the time to shower me with “unctuous praise” via Skype one week before submission, demonstrating once again that despite the distance he is one of my best friends; Melanie and I shared numerous enlightening conversations starting with a ‘Hey girl’ and finishing in a hip cafe in Amsterdam.

I’ve been lucky to have the likes of “sound” Olly Spinney, Andy Kirton, Benni Magro, Carlo Rainieri, Fred Horton, Joey Montgomery and Andries De Jong to share the PhD life in Manchester. Marta Cabrera and Jansan Favazzo came over briefly and made everything infinitely better.

More importantly, this wouldn’t have been a period of mostly sustained sanity if it wasn’t for the women who have stood by me during my thesis.

In the Manchester Philosophy department, Aby Connor has been my companion and role model, constantly deserving the label of ‘superwoman’: kind, hard-working, attentive and strong. Thanks, Aby, for ignoring me when I said I was not a hugger. Lydia Farina has been a calm, reflexive counter-point (except when there was Greek dancing), taught me how to grow up in style without losing the fun offered unlimited cheering up –together with her partner Mema Evangelatou and their drama-dog Byron.

In the Manchester non-philosophy world, Rita Cervetto showed up one day in my sofa and became one of the pillars of my life, always being at the other side of the line despite my phobia of answering phones. Elia Giménez taught me to love myself and to ask for help.

In Edinburgh, I left three women behind who have been a constant example in my talks as ‘my romantic friends’. Naomi Swainson, Carmen Galiano and Anastasia Kakokariti: I honestly can’t convey here how important you are and have been these four years, for my ‘pehaichdé’ and my life. All I can do is write is a cheesy ‘I love you’ for posterity and send you one million pollos to summarise the ten pages of gratefulness you each deserve.

Scattered through several countries in Europe during this period despite being only two people, Marta Massé and Iwona Małecka’s S-team have been the best people to organise summer gatherings with (even ones including long hikes across cliffs).

My Andalusian women: Marta, Cary, Fátima, María José, Mercedes, Rosa, María E., María M., María B., Inma R., Inma G., Noemí, Paloma, Ana, Clara, and Enca. You have made me laugh when I was down, listened to boring academia chat and checked
on me all the time, while inviting me to your weddings and taking me out to drink *verdejo* and *vino dulce* – a fun holiday distraction away from philosophy books.

Above all, I thank my parents, Pilar Cantero and Manuel López, for making this possible. First, with their generosity, paying for my master’s fees together with my aunt, Pepa Cantero, and my dad’s partner, Lola Niño. More importantly, my parents made me the kind of person that would think it was a good idea to do a PhD on love in the first place. Thanks, mum, for teaching me that I should live my life as I wish even if it breaks people’s expectations; and thanks, dad, for teaching me that in any project that I start I can, and should, aim for the highest. Thanks to you, and to my siblings Angela, Víctor and Carmen, for your unfaltering faith in me.

Last but absolutely not least, thanks to Jack Casey for doing all of the above (except maybe the paying for stuff part). Some of my ‘fluffy’ ideas have only become serious arguments after running them through him – Jack even found the Pete Doherty interview that ended being one of the foundations of my argument. More importantly, I would have spent most of the past four years swarmed in academic chat if it hadn’t been for our long conversations about politics, love, wokeness, dangerous border crossings, Spanish and British idiosyncrasies and, to my dismay, grounding and football – which made him lose the bus to Mordor more than once. Thank you, Jack, for being my best friend.
Dedication

This thesis is dedicated to the memory of my grandma, Amparo Jiménez Caño, who would have proudly cried over the phone because of it while checking the international weather news –and who taught me unconditional love.
Introduction

It is a common trope of philosophical analyses of emotions or everyday practices to brand themselves as the study of ‘underinvestigated phenomena’. It is certainly a frequent move in the philosophy of love, which represents an oversight of the fact that love has been investigated by many of the ‘greats’ in the History of Philosophy: from Plato and Aristotle to Aquinas and Augustine; Jean-Jacques Rousseau, Friedrich Nietzsche or Arthur Schopenhauer; Simone De Beauvoir and Jean Paul Sartre. To label love as ‘underinvestigated’ may seem to make certain sense within the context of Western analytic philosophy, seemingly more concerned with ontological puzzles, the functioning of the mind, or questions regarding language. Still, love has been an object of conceptual analysis for prominent analytic philosophers such as Martha Nussbaum, Harry Frankfurt, David Velleman, Robert Nozick or Ronald de Sousa. Currently, love is a growing topic in several fields, mainly philosophy of emotion, moral psychology and practical philosophy. It only makes sense: besides more traditional philosophical questions, the analytical method has a lot to contribute to the understanding of features of people’s everyday lives – like love.

However, there is still something to say about the status of the topic of this thesis – being in love – in analytical philosophy. Being in love can be defined as a feature of romantic love. Romantic love, which broadly is the love people have for their partners (or the people who they would want to be their partners, all things considered) is not underinvestigated. But it has been investigated in an odd way.

Frankfurt, who thinks that love is the wholehearted endorsement of what one cares about and hence a necessary component of an authentically lived life, says that relationships that are primarily romantic or sexual do not provide very authentic or illuminating paradigms of love as I am construing it. Relationships of those kinds typically include a number of vividly distracting elements, which do not belong to the essential nature of love as a mode of disinterested concern, but that are so confusing that they make it nearly impossible for anyone to be clear about just what is going on. (Frankfurt, 2004: 43)
For Velleman, love is an appraisal of the loved person’s incomparable value (1999: 365). This appraisal, however, does not happen in “blind, romantic love” (Velleman, 1999: 351):

[What I have in mind is the love between close adult friends and relations—including spouses and other life-partners, insofar as their love has outgrown the effects of overvaluation and transference. (Velleman, 1999: 351)

These suspicions about romantic love are shared even by philosophers who do not explicitly reject the value of romantic love. Robert Nozick offers an account specifically about romantic love, with a caveat. For Nozick, romantic love has two stages. There is a first stage of being in love which makes a person absorbed in the object of love to the point of her other concerns and responsibilities becoming “a minor background detail” (Nozick, 1989: 418). The second, which he calls romantic love, requires a switch from being in love: “if given the opportunity, [being in love] transforms itself into continuing romantic love or else it disappears” (Nozick, 1989: 418). Although Nozick does not explicitly make a value judgement on first stage love like Velleman above, his focus on second-stage love as the worthy object of analysis reveals that, in his account, being in love is second-rate love—actually, valuable romantic love requires ‘getting over’ being in love.

This is odd, for two reasons. The first reason this portrayal of romantic love that it is very far removed from people’s experiences and conceptions of it. Contra Frankfurt, people seem to attach authenticity to romantic love; we want our love to be real; for many having real romantic love in their lives seems to be one component of authentic lives. Contra Velleman and Nozick, being in love seems to be an essential feature of romantic love, which is not understood by most merely as obsessive infatuation at the beginning of a relationship. When one hears the phrase “I’m not in love with you” one hardly expects that what follows is a more stable, more valuable stage of the relationship. On the contrary, it tends to mean that the relationship is over or that it will never happen. People want to be in love with their partners, and their partners to be in love with them. Being in love is constitutive of romantic love, not a preliminary phase that needs to be overcome to reach valuable love.
The second reason is that conceptual analysis should not be value-laden: it does not seem like the best philosophical practice to aim to explain a phenomenon –love– and discard a very extended instance of that phenomenon –being in love– as not valuable without further discussion. Furthermore, conceptual analysts should not refuse to define certain phenomena without proper justification: Frankfurt’s claim that “it is nearly impossible for nearly anyone to be clear about just what is going on” (1998:43) in romantic love seems like a philosophical cop-out.

This oddity has long loomed over the philosophy of love. Frankfurt, Velleman and Nozick —and many subsequent accounts of love influenced by them— have undoubtedly made a valuable contribution to the understanding of certain features of love. Some of the claims of this thesis rest on their work. However, they are still found wanting in the analysis of being in love, in virtue of the underlying value judgements on romantic love. A great proportion of accounts on love draw —at least partly— from these three authors without addressing their stance on romantic love, while implicitly or explicitly including romantic love as their explanandum (there are some exceptions, notably Foster, 2008; McKeever, 2019). Only recently have philosophers explicitly aimed to overcome the downgrading of romantic love, which I think Gary Foster rightly attributes to a Christian and Kantian presupposition that expecting love to be reciprocated somehow “cheapens” love (2008: 242). Much of the literature on love I draw from to build my own claim has been published in the last decade; often during the process of writing.

Still, there is very little philosophical literature specifically on being in love. I believe this stems from the way the debate has been framed. Taking love for an attitude and following the analytical tendency to explain attitudes in terms of reasons, the debate has been skewed towards the rationality of love. Discerning whether romantic love can be analysed in terms of reasons and, if so, establishing those reasons, is indeed a worthwhile project. Maureen Sie says that ‘stage one’ romantic love “makes people act immorally, steal, lie, kill, and even lead whole armies astray” (2018: 41). Although this is a tongue-in-cheek comment and Sie is just describing a specific stage of romantic love called ‘limerence’, it is quite illustrative of a generalised conception of romantic love as irrational. So, it does make sense to ask which are the reasons of love. Within the rationality-oriented approach, the big question has often been which properties of the loved person rationally justify love. Usually, rationality-oriented
philosophers of love have been divided between those who think love is grounded in the loved person’s properties (Delaney, 1996; Keller, 2000), those who think that love is rationally justified by the relation between the lovers (Kolodny 2003), or those who think it is a mix of the two (Naar argues in his (Naar, 2017) for subject-relative reasons for love, which are the loved person’s properties expressed in a relationship). In turn, this has resulted in a focus on the so-called question of fungibility or irreplaceability: is it rational to ‘trade-up’ your beloved for someone with better qualities? Philosophers have devoted their efforts mostly to defend that it is not (Grau, 2010; Naar, 2019). The answer to rationality-oriented philosophy of love has largely rested on the claim that love is not justifiable in terms of reasons, i.e. these are the wrong questions to ask (that’s Frankfurt’s position in his (Frankfurt, 2004); see also Thomas, 1991; Zangwill, 2013). Still, this answer has kept the debate on love within the realm of rationality. To quote Ulrika Carlsson, a recent supporter of the no-reason view, “[t]he first mistake made by the intellectualist view on love was to assume that “rational” was the greatest honorific a philosophical theory could bestow upon love” (Carlsson, 2018: 1408). I share Carlsson’s sentiment, but in my view, focusing on defending a no-reason view on love continues to honour the rationality approach in virtue of ignoring other interesting questions, thus perpetuating the odd status of romantic love in philosophy. There are numerous arguments on the grounds for being in love (let them be rationally assessable or not), but we are none the wiser as to what being in love is.

This thesis aims to avoid these problems, and to solve the conceptual gap on being in love. It is a descriptive project, not a normative project understood in the sense above. Normatively describing a phenomenon in terms of its rationality and/or its value is a valid and worthwhile metaphysical approach, but it is not the project I have in hand. In what follows, I offer an account of what being in love is. It is neither grounded in the value nor in the reasons of being in love. This approach is philosophically justified: it follows the belief that before defining reasons or value of a phenomenon, we need to describe the phenomenon first. Mine is not a description of romantic love in general, since I acknowledge that there is more to romantic love than being in love. But it rests on the conviction that an account of romantic love which rejects or ignores being in love as worthy of conceptual analysis is, at the very least, odd; and that defining being in love is, in part, defining romantic love.
Chapter 1

Love, caring, and identity

It seems intuitive to say that being in love with someone is, or at least entails, caring about that person. In this chapter, I describe, and reject, the approach that I call the ‘caring view’, which is based precisely on that intuition. Eventually, I propose a change of direction in order to define what being in love is. Instead of focusing on the loved person, I argue, being in love may be best explained by focusing on the lover.

I begin in §1.1 by presenting Frankfurt’s (1998, 2004) view: caring about someone is for their interests to become analogous to your interests. Bennett Helm (2010a, 2010b) criticises Frankfurt’s approach and says, instead, that caring about someone is having a specific type of concern about their identity. For Helm, caring entails being rationally committed to a pattern of emotions towards the loved person’s identity.

Given that Helm’s view is based on the loved person’s identity, I present his account of identity in §1.2. According to Helm, a person’s identity is composed by her conception of a life worth living and her values, which are interdependent. Unlike traditional accounts of personal identity, Helm accepts that a conception of a life worth living or values need not conform to universal or moral standards of well-being.

In §1.3, I present several objections to Helm’s view, given that it does not accommodate non-harmonious love. Firstly, Helm cannot accommodate cases of ‘love in spite of’ problematic values (which do not conform to universal and/or moral standards of well-being) or opposing values. Secondly, Helm cannot accommodate cases of neutral values where the lover does not care about the loved person’s values.

Finally, in §1.4 I introduce an alternative approach which incorporates what the caring view gets right about love, while shifting the focus of analysis from the loved
person to the lover. Instead of defining love as a way of caring for the loved person, I propose to explore the change that being in love constitutes for the lover.¹

1.1. The basics of caring

1.1.1. Frankfurt on love and identification

Harry Frankfurt has one of the most influential accounts of love as caring. For Frankfurt, love has “four main conceptually necessary features” (2004: 79): disinterested concern for the well-being of the loved person, focus on a specific person, identification with the loved person and love not being a product of choice. Here, I focus solely on the necessary feature of the identification of the lover with the loved person.² Frankfurt states that a person who cares about something identifies himself with what he cares about in the sense that he makes himself vulnerable to losses and susceptible to benefits depending upon whether what he cares about is diminished or enhanced. Thus he concerns himself with what concerns it, giving particular attention to such things and directing his behaviour accordingly. (Frankfurt, 1988: 83)

Then, what a person cares about i) affects a person’s well-being (one becomes “vulnerable to losses and susceptible to benefits”); ii) directs a person’s attention and iii) directs a person’s actions. This seems like a plausible route to explain what being in love is, as the following example shows. Manuela is in love with Casey. With Frankfurt’s account in hand, in virtue of being in love with Casey, Manuela’s well-being is affected. Manuela suffers when Casey suffers (she is vulnerable to his losses) and flourishes when Casey flourishes (she is susceptible to his benefits). If Casey

¹ Although ‘lover/loved person’ may sound somewhat clumsy at first, I think it is better that the commonly used ‘lover/beloved’—which, like Amélie Rorty, I find “quaint” (2016: 344). I use couples in my examples, but I assume that what I say can potentially apply to both monogamous and polyamorous love. Those who are convinced by arguments against polyamory can interpret my thesis as pertaining to monogamous romantic love only. For an argument against polygamy (a form of polyamory) see Brooks (2009). For a defence of polyamory, see Jenkins (2017), McKeever (2017) and Brunning (2018).

² As I showed in the introduction, Frankfurt explicitly rejects that romantic love is among the “authentic or illuminating paradigms of love” (2004: 43). I have already explained that downgrading romantic love as a sub-class of love is something that analytic philosophers are rightly overcoming, so I just assume that romantic love can also be defined by these four necessary features. For an account of Frankfurtian romantic love which avoids the contradictions between downgrading of romantic love and these four necessary features, see McKeever (2019). For a rejection of Frankfurt’s four features of love, see Soble (2011).
loses his job, Manuela not only feels sad for him or worried about how they are going to pay the bills—her own well-being diminishes because Casey’s well-being has diminished. If Casey gets a promotion, Manuela not only feels happy for him or enthusiastic about the holiday in the Maldives they can finally afford—her own well-being increases because Casey’s well-being has increased. Manuela concerns herself with what concerns Casey (like, for example, his career), gives particular attention to things related to Casey (she thinks all night about whether he will get the promotion), and acts accordingly (she supports him through job-seeking; she plans a promotion party). Casey’s career is not just something Manuela pays attention to because it is among Casey’s interests. Casey’s career is, in virtue of loving Casey, one of her interests:

The lover is invested in his beloved: he profits by its successes, and its failures cause him to suffer. To the extent that he invests himself in what he loves, and in that way identifies with it, its interests are identical with his own. It is hardly surprising, then, that for the lovers selflessness and self-interest coincide. (Frankfurt, 2004: 61–62; emphasis in original)

The idea that the loved person’s interests become the lover’s interests is the basic foundation for many theories of love. This is the case of the so-called ‘union accounts’, which define love as a sort of merging of the lovers’ identities. The idea of merging has been cashed out with differences of degree, from Robert Solomon’s ‘fusion’ of the lovers’ identities into one (1988: 24) to Robert Nozick’s desire to form a ‘we’ as a new separate identity from the lovers’ (1989: 70)–Nozick’s weaker union view has been mirrored by Neil Delaney (1996) and Marilyn Friedman (1998). Frankfurt’s is not a union view: he does not understand the loved person’s interests becoming identical to the loved person’s interests as a merging of identities. On Frankfurt’s account, love is volitional, since loving someone entails a constraint of one’s will (his is what is usually known as a ‘robust concern’ view).3 Without the need to get into volitional claims, however, we can still discuss the idea of the loved person’s interests becoming identical to the lover’s interests, which is common to Frankfurt and the union views. Bennett Helm summarises that idea as follows:

3 Other robust concern views are Soble (1990), LaFollette (1996), and White (2001). See Smith (2011) for an account that combines Frankfurt’s robust concern with Nozick’s union view.
Your concern for the identity of the other is not merely analogous to your concern for your own identity: it is a part of it. In identifying with your beloved, you make her cares and concerns, her interests and values, become a part of your identity, so that you care about her as a part of caring about yourself and thereby tie your own well-being to hers … Intimacy, therefore, requires incorporating her well-being into your own. (Helm, 2010: 10; emphases in original)

For Helm, to care about something is for that thing to have import in you, which means to have concern about its well-being (2010b: 24). This concern implies that the object of caring is “worthy of attention and action on its behalf” (2010b: 24). Caring then requires the carer to be concerned about the well-being of the object of her caring, and to be prepared to act toward the promotion of that well-being. In that sense, Helm’s shares Frankfurt’s basic notion of identification as grounded in well-being, attention and action. However, as I show next, he rejects Frankfurt’s view.

1.1.2. Helm’s critique

Targeting both union and robust concern views, Helm complains that they suffer from what he calls a “tendency toward individualism” (Helm, 2010a: 9). Individualism, for Helm, comes in two forms: an “egocentric conception of intimate concern” and an “individualist conception of autonomy” (Helm, 2010a: 9). I will come back to Helm’s claims on the individualist conception of autonomy in §1.4, but with regard to Frankfurt, it is his claim on egocentrism that is most relevant.

For Helm, the reason why traditional views of love—including Frankfurt’s—are egocentric is because they require the loved person’s identity to become part of the lover’s identity. Helm says that concern should not be understood as intimate in virtue of its contribution to one’s identity. Instead, Helm argues that we should come to understand what is distinctively intimate about love in terms of a distinctive kind of concern for the identity of another as the particular person he is, a concern that it is the same in kind as the concern you have for your own identity but without presupposing that
you thereby make your beloved’s interests and identity a part of your own. (Helm, 2010a: 18)

This is as much a critique of union views as it is of robust concern views. According to Helm, both fail to account for the intimate character of concern in relationships like romantic love or close friendships. Going back to the example I gave above, if it were not because Casey’s identity had become part of Manuela’s identity, Manuela’s concern would not be intelligible as intimate. In that picture, Manuela’s concern is intimate because Casey’s interests are now hers. This happens in union views, where Manuela’s and Casey’s identities are incorporated to each other’s in one way or another, and in robust concern views, where Casey’s well-being is incorporated into Manuela’s—and vice versa.

For Helm, that way of understanding concern fails to account for what makes caring for persons who are one’s lover or friend as intimate. On Helm’s account, to say that the lover has to be invested in the loved person is egoistic, given that such a claim grounds concern for the loved person in the lover’s own interests—even if it is an interest for the loved person. The objection is that if the loved person’s interests become the lover’s interests, then the lover is merely concerned, attentive and active with respect to her own interests. Helm argues that intimate identification requires the lover to be concerned, attentive and active for the loved person’s interests, which are not the lover’s own. Helm would then consider that Manuela’s concern for Casey is intimate because Manuela cares about Casey’s interests in virtue of being his interests. It is the fact that it is his interests that we can say that Manuela cares for Casey for his sake. I give more details of Helm’s view in the next section—while, for the moment, remaining neutral with respect to Helm’s critique.

1.2. Helm’s intimate identification

1.2.1. Caring about persons

Like Frankfurt, Helm believes that caring about someone is to be concerned, attentive and active about what that person cares about. Unlike Frankfurt, Helm does not believe that the lover incorporates the loved person’s interests to her own set of interests. Instead, Helm highlights the idea that caring about a person is having
concern for her identity (“concern for the identity of another as the particular person he is”, Helm 2010a: 18). He calls this type of caring ‘intimate identification’.

What is to be concerned about someone’s identity? Helm says that the well-being of a person is partly defined by her personal values, which constitute “her sense of what contributes to the kind of life worth her living” (Helm, 2010a: 42). This sense of a life worth living “constitutes one’s identity as this particular person” (Helm, 2010a: 42; emphasis in original). In intimate identification “you come to have a concern for your beloved’s identity that is the same in kind as your concern for your own” (Helm, 2010a: 43). Hence, to care about a person means to care about her identity understood in that way. Helm draws from the philosophy of emotion to explain how it is possible to care about someone’s values without the need to incorporate them to one’s own.

Briefly, in his account of emotions Helm distinguishes between the target or intentional object of an emotion (what the emotion is about) and the formal object (the evaluation which constitutes the emotion). He gives the following example. If I am angry that rabbits are eating the tomatoes in my garden, the target of anger are the rabbits, while the formal object is the evaluation of their actions as offensive or enraged. Helm adds a third element, the focus of the emotion: “the background object whose import to the subject makes intelligible the evaluation of the target in light of the formal object” (2010b:25). That is, the thing that the subject cares about: the tomatoes. If I did not care about my tomatoes, I would not find the rabbits offensive, he argues.

Caring about tomatoes entails that I will feel certain emotions regarding the tomatoes: “fear that the groundhogs will dig under the fence, worry that the cold, wet spring will rot their roots, excitement at the tomatoes finally start ripening” (Helm, 2010b: 25). Helm says that caring entails a rational commitment to feeling these emotions, since “in feeling an emotion one is in effect committing oneself to the import of its focus” (Helm, 2010b: 25). This emotional commitment results, over time, in a pattern of emotions to which we are rationally committed. In this case, caring about tomatoes commits a person to a pattern of emotions focused on the

---

4 Helm explicitly rejects that his view on personal identity has any bearing on personal continuity, i.e. what makes a person at t2 the same person at t3 (2010a: 136, fn. 52). I briefly come back to this when I separate my thesis from that question in §2.1.
tomatoes—elation when they start growing, anger when the rabbits eat them, worry when frost threatens the harvest, etcetera.

For Helm, intimate identification follows this same structure. However, intimate identification is a “deeper” mode of caring than caring about tomatoes, given that it is caring about a person’s identity (2010b: 27). While emotions towards objects like tomatoes have one focus (the thing that the person cares about), what Helm calls ‘person-focused’ emotions (2010b: 28) has two levels of focus. The focus of the emotion is the particular person which is the object of import, while the subfocus of the emotion is the particular value of that person’s identity.

Distinguishing between focus and subfocus allows Helm to explain identification without the need for the loved person’s interests to become one’s own. Helm gives the example of his feelings of pride when his wife wins a bagpiping competition (I assume, although he does not say explicitly, that bagpiping is one of his wife’s values). He cares about his wife, and that grounds his own pride at his wife’s winning. His wife is, then, the focus of the emotion, while bagpiping is the subfocus. Bagpiping is the specific value that Helm cares about in virtue of it being owned by his wife, about whom he cares and who is the focus of the emotion. Helm does not have bagpiping as one of his own values. Bagpiping is something he cares about because it is a value of a person he cares about. In virtue of loving his wife, Helm is rationally committed to a pattern of emotions (including pride) that have his wife as focus and his wife’s values as subfocus (for example, bagpiping). In that way, Helm offers an alternative to the Frankfuritian approach, where the loved person’s interests become one’s own.

Helm’s critique can be accepted without the need to accept or reject his qualification of Frankfurt’s view as egoistic. Regardless of whether one agrees with Helm’s egoism objection, his view is better suited to explain being in love than Frankfurt’s, for two reasons.

First, Frankfurt does not really elaborate on the kind of interests of the loved person that become interests of the lover. If being in love entails concern about the loved person’s interests, surely it is not required that the lover is concerned for all the loved person’s interests. We may want to say that being in love requires that one cares about one’s wife’s bagpiping prize, which is partly what makes her life worth living. But surely one does not need to share anger when one’s husband is angry about
rabbits eating his tomatoes. Frankfurt’s requirement that the loved person’s interests are ‘identical’ with the lover’s interests is so general that it would require that one’s well-being diminishes when one’s husband suffers regarding his tomatoes being eaten by rabbits. That is just too stringent for being a plausible explanation of what being in love is.

Secondly, even if the diminishing of the loved person’s well-being causes the lover to suffer too – even in cases when one suffers when one’s husband’s tomatoes are eaten by rabbits –, that does not mean that the loved person’s values have to become the lover’s values. For example, Aby might be a long-time Manchester United fan; being a supporter of the club may well be one of her values – part of the life she considers worth living. In Helm’s picture, when Aby is happy because Manchester United wins a match, her partner Adam being happy is not grounded in Manchester United being one of his values. The faring of Manchester United may not have any influence on whether Adam thinks he is living a life worth living. Adam feels happy because the club is one of Aby’s values, so in virtue of caring about Aby, he is rationally committed to a pattern of emotions towards Aby with Manchester United as a subfocus.

Helm’s view is better equipped than Frankfurt’s to explain what being in love is, given that it is quite frequent that the lovers’ interests do not align completely. It does not seem necessary that we have the same values, understood in Helm’s terms, as the people we love. Any academic philosopher with a non-academic partner can easily understand this – often partners actually cared very little about philosophy before they met us, and only care about it because we do.

In the context of Helm’s view, being in love would then be a rational commitment to a pattern of emotions which have the loved person as focus, and the loved person’s values as subfocus. This pattern, which is bi-directional (Aby has it for Adam and Adam for Aby), results in a change in the lovers’ identities, who end sharing a “single evaluative perspective” (Helm, 2010a: 260) in virtue of both being rationally committed to patterns of emotions towards certain subfocuses (their respective values). To explore whether this is a plausible approach to explain what being in love is, I offer more details on Helm’s view on identity.
However, as I argue next, this view is also problematic, given that it does not account for non-idealised cases of being in love. In order to expose the problem, first I need to give further details on Helm’s conception of identity.

1.2.2. Identity, well-being and a life worth living

According to Helm, intimate identification is deeper than mere ‘caring’ because it consists in caring about the loved person’s identity, which is analogous to the concern one has towards one’s own identity. To care about someone’s identity (one’s own or others) is to care about their values: what makes their life worth living.

The idea that a person’s identity is constituted by the kind of things she considers worth her living has also been famously defended by Christine Korsgaard. Korsgaard claims that a person’s identity is defined by her actions. Through our actions, we endorse what she calls ‘practical identities’: “descriptions[s] under which you value yourself and find your life worth living and your actions worth undertaking” (Korsgaard, 1996: 101). These descriptions of ourselves give us reasons for action. Our actions are then justified by reasons which we source from the type of life we consider worth living. The idea of a life worth living is grounded in importance: “It may be important to you that you are a human being, a woman… someone’s lover” (Korsgaard, 2009: 20). For example, I have the practical identity of being a philosopher, and my friend Mercedes has the practical identity of being a mother. Mercedes is also an architect, but that role is not a source for reasons in the same way that being a philosopher is a source of reasons for me. Surely, being an architect gives Mercedes reasons to act in certain ways, like going to work on time, drawing plans and updating her knowledge on Revit. But it is not part of her idea of the type of life she considers worth her living like being a philosopher is for me. The difference is that for me, the role I acquire with my profession is important, while for Mercedes hers is not.

Helm rejects Korsgaard’s view. He argues that if a person’s identity is constituted by her values, then valuing is conceptually prior to a person’s identity (Helm, 2010a: 120).

My claim is that loving yourself means finding yourself to have import as this person, as a person with a particular identity as such, for without
this it is unclear what could be meant by finding yourself to have import “for your own sake”: your “sake,” your well-being, is defined by your identity as the person you are. Consequently, loving yourself presupposes that you have a determinate identity as this person. (Helm, 2010a: 130; emphasis in original)

Unless values and the conception of a life worth are interdependent, but separate, Helm thinks that we cannot account for the fact that people’s identities are malleable. He makes an illuminating analogy:

In some respects, therefore, a person’s identity is like a building that is constantly undergoing remodelling, with parts being demolished, others being added, still others being modified, and all without a fixed, determined blueprint. (Helm, 2010a: 132)

Determining that conception is the expression of a person’s autonomy:

A person… is a creature that has the capacity to have a say in defining its own identity as such and so in determining its well-being as the particular person it is; this is our autonomy, our capacity for self-determination. (Helm, 2010a: 131; emphasis in original)

I do not stop here to discuss whether Helm is right or not on whether practical identities are actually isolated from each other in Korsgaard’s view. It suffices for my purposes to say that since Helm discusses this interdependence explicitly and Korsgaard does not, I work from Helm’s view on identity and adopt his term ‘values’ instead of ‘practical identities’. The reason is that the arguments I will make later rest on the possibility of values and a conception of a life worth living sometimes coming apart from each other.

In this way, we have a clearer picture of what values and the conception of a life worth living mean for Helm. Helm says that one’s well-being is defined by one’s conception of a life worth living and one’s values, which are interdependent but separate. Helm specifically says that values need not conform to objective standards:

It should be clear that this notion of valuing is distinctively personal insofar as it is both relative to the individual person and definitive of
who she is as a person. Such *personal values*, as we might more properly call them, are distinct from moral or other universal values. In particular, although I might recognize that certain works of art or nature scenes have value, I need not (though I may) personally value these things by finding somehow to be a part of the kind of life worth living. Moreover, it should be clear that such personal values are a matter of evaluative attitudes a person in fact has, as distinct from what she should value. (Helm, 2010a: 98; emphasis in original)

This is a promising view on personal identity, given that it is true that people value things which are distinct from ‘what they should value’. Some people care about art and philosophy, but some others care about being rich or about being members of an organised crime syndicate. For both the philosopher and the mafioso, these are things that make their lives worth living. A theory of love as caring must accommodate caring for disvaluable things. The truth is that we do not only love valuable people, if ‘valuable’ is to be understood according to moral or universal standards. Like Samuel Scheffler points out, “a relationship may be unhealthy or exploitative or oppressive”, and the fact that a person values it does not mean that it is worth valuing (2010: 105). I come back to this in §1.3.1, and give an extended argument in favour of disconnecting romantic love from morally right action in §4.1.3. For the time being this suffices to give Helm’s view on identity the upper hand.

To explain what being in love is following Helm’s account, we need to start with the loved person’s values. Marta is a philosopher, and being a philosopher is one of her values. Being a philosopher constitutes part of her identity in relation to other values she has and her conception of a life worth living (‘a life aimed at understanding the world’). Marta’s promoting of her own well-being is a promotion of her identity, realised partly as the promotion of her value of philosophy (partly, because she also needs to promote other values she has). For Janis to love Marta, in Helm’s view, is for Janis to value Marta. That means that Janis cares about Marta for her sake, i.e. for her identity as the person she is. That also means that Janis has philosophy as a subfocus in virtue of caring about Marta. Marta’s identity is the focus of Janis’s caring, and her being a philosopher is the subfocus of Janis’s caring. Janis’s being in
love with Marta commits him to a pattern of emotions towards Marta’s identity and values in this sense.

Throughout his discussion of values, Helm keeps giving examples which can all, more or less, be accepted as potential components of a life worth living: being a professor (2010a: 118), being a father (2010a: 122), and, as seen above, playing the bagpipes. It seems intuitive to say that being in love with a person entails caring for values which constitute their identity such as being a professor, being a father, being a musician or, like in my example, being a philosopher. But let us remember that, for Helm, personal values are not necessarily in accordance with universal or moral values. All these examples easily conform to universal or moral values. How does Helm’s theory of caring fare when the loved person’s values do not conform to universal or moral standards?

1.2.3. Problematic identities

English musician Pete Doherty, former frontman of the rock band The Libertines, has experienced decades-long drug addiction, widely documented in the media. In an interview, Doherty states the following:

For a long time … I didn’t use any drugs. On and off over the years I haven’t, but it’s become such an intrinsic part of my existence. It’s like talking about a brother, really. I can’t really separate myself from them. They’ll always be there. (Doherty, 2019: n.p.)

Doherty’s account departs from commonplace first-personal accounts of drug addiction, where drugs are something undesirable that need to be overcome. On the contrary, for Doherty, talking about drugs is like “talking about a brother”; drugs are an “intrinsic” part of his existence. Doherty goes on to say that drugs are one of his “favourite conversations”:

I’m fascinated by drugs, always have been. How they are controlled, how they work economically, socially. How they work metabolically [sic] and metaphysically and intellectually. And I love to be around them, I’m fascinated by the clothes, by the culture. (Doherty, 2019: n.p.)
It seems that Doherty cares about drugs. If we go back to the basic definition of caring, caring is grounded in well-being, attention and action. It is indisputable that Doherty’s caring about drugs fulfils the last criteria in the same way Marta’s caring about philosophy. Doherty is disposed to attend to drug-related themes in the same way Marta is disposed to attend to philosophy-related themes. Doherty acts accordingly, looking for drug-related themes which he enjoys discussing, like Marta enjoys discussing philosophy (and predictably looking to engage with drugs by taking them, like Marta looks to engage with philosophy by writing it). Prima facie, the difference between Marta and Doherty is that it may be said that drugs are detrimental to well-being objectively, while philosophy is well-being to well-being objectively.

Doherty is a paradigmatic case of what here I call ‘problematic identities’. Korsgaard, for example, would plausibly argue that Doherty lacks integrity as an agent, given that he is not acting according to universal reasons (reasons to not engage in self-destructive behaviour). Frankfurt, in fact, has an example which almost exactly mirrors Doherty: the Willing Addict, who “would not have things any other way” and thus is morally responsible for his desires (Frankfurt, 1971: 19). I call ‘problematic identities’ the identities which do not conform to moral or universal standards. With Helm’s theory in hand, however, we can understand problematic identities leaving these moral considerations aside.

In Helm’s view, we cannot say that Doherty’s caring about drugs is necessarily damaging to well-being, even when a universal or a moral standard would reject drugs as conducive to well-being. Recall that, for Helm, promoting one’s well-being is promoting one’s identity understood as one’s values and one’s conception of a life worth living. Doherty’s conception of a life worth living could plausibly be summarised as ‘a life of sex, drugs and rock and roll’. Valuing drugs is in accordance with Doherty’s conception of a life worth living; like being a philosopher is in accordance with Marta’s conception of a life worth living (as ‘a life aimed at understanding the world’). Doherty clearly has drugs as one of his values, an expression of ‘a life of sex, drugs, and rock and roll’.

However, despite having drugs as one of his values and a conception of a life worth living as a life of sex, drugs and rock and roll, Doherty knows that there is something
wrong with his identity. In another recent article about Doherty, the writer reports the following exchange:

Would you like to be clean? “Yes, a part of me would. Just so I can feel things. There are so many people in my life who deserve better. It really is a mental deficiency.” Would you be more productive if you were drug-free? “I’d be a force to be reckoned with! I’d have money and self-respect and clean hands.” His fingers are filthy. (Hattenstone, 2019: n.p.)

Doherty says he lacks self-respect. Helm acknowledges that this happens when personal values which are not in accordance with universal or moral standards:

[W]e might understand self-hate to be the sense that your current values, your current sense of the kind of life worth your living, is fundamentally misguided so that, even though you may be living up to these current values, you nonetheless are not living as you ought. In short, understood this way self-hate is the sense that your identity as you now understand it is somehow rotten. (Helm, 2010a: 140)\(^5\)

It is, then, possible to have problematic values and a problematic conception of a life worth living which are in conflict with the values and conception and a life worth living that you ought to have. Doherty’s idea of the general conception of a life worth living partly differs from the conception that he ought to have. It is important to insist on partly because only “a part of him” wants to be drug-free; he explicitly says that he is fascinated by drugs and loves to be around them. It is only partly that his conception of a life worth living comes apart from his valuing taking drugs. But, all in all, there is a dissonance between values and identity in Doherty’s case which is an example of understanding one’s identity as ‘somehow rotten’.

It could be argued that Doherty is immoral, ambivalent, or that Doherty’s free will is compromised due to his first-order desires. But there is no need to get into those discussions, given that, as I show next, problematic identities put Helm’s views on caring about oneself and caring about other people in conflict with each other.

---

\(^5\) Helm thinks that self-hate is ultimately motivated by self-love: is it out of caring for your own identity that you recognise your identity as flawed. Self-hate is, for Helm, ‘deficient self-love’ because it fails to motivate you to improve your identity (Helm, 2010a: 140).
1.3. Rejecting the caring approach

I now present an objection to Helm’s account of intimate identification. I argue that some people are loved in spite of their values, not because of them. I then argue that, in general, the caring view does not accommodate non-idealised, but nonetheless very common, instances of romantic love. On the basis of these objections, I argue that the caring approach is unsuited to explain what being in love is.

1.3.1. Love in spite of problematic identities

As I explained in §1.2, Helm defines love in terms of intimate identification: a rational commitment to a pattern of emotions which have the loved person as focus (i.e. his wife), and the loved person’s values as subfocus (i.e. bagpiping). Specifically, the lover is rationally committed to the same emotions that the loved person has towards her own values:

To feel this pride is to be committed to the import she has as this person and so to the import bagpipes have as a part of her identity as such. Such commitments are, therefore, to feeling other person-focused felt evaluations with the same focus. Indeed, these will be person-focused felt evaluations I ought to share with her: when she is proud, ashamed, anxious, self-assured, and so on, where these are emotions focused on herself, I ought similarly to feel pride, shame, and so on for her sake. (Helm, 2010a: 156)

This seems immediately wrong in the case of people with problematic identities. In the example of the bagpiping prize, Helm feels pride which is focused on his wife and subfocused on bagpiping, in virtue of bagpiping being one of his wife’s values. It would be slightly bizarre to claim that, equally, someone who loves Doherty is rationally committed to feeling happy when he is happily taking drugs. Let us remember that Helm accepts that people can have problematic values, given that values do not need to conform with moral or universal standards. So, in theory, problematic values –like drugs– should rationally commit a person to the same patterns of emotions as non-problematic values –like bagpiping. It follows that, indeed, Helm’s account commits him to the bizarre claim: Doherty’s partner seems to be rationally committed to happiness when Doherty happily takes drugs. However, Helm anticipates this scenario and refutes the bizarre claim:
I ought not to be proud of him for upholding values I think are generally abhorrent; nor could I more generally share this value with him, even if it is for his sake. For to do so would be for me to be inadequately sensitive to the value these things intelligibly can have (by my lights), and I would thereby fail in my commitment to his well-being. (Helm, 2010a: 159)

Helm continues with a claim that I believe stands in conflict with the rest of his account:

[I]f my beloved is about to embark on a course of action I find prohibited not just for myself but for anyone, or if he fails to value something I think he must, then, other things being equal, I ought to object out of a concern for his well-being. For in such a case simply to defer to his view would be, according to me, harmful to his well-being and so contrary to my concern for him; indeed, if he continues in this way, then I ought to be ashamed of him for failing to live as he ought in this respect, regardless of his failure to be ashamed of himself. (Helm, 2010a: 160)

Here Helm seems to brush aside his previous definition of well-being. Promoting a person’s well-being, let us remember, is to promote what they conceive as a life worth living. Helm explicitly rejects the need for values to conform with universal or moral standards. So even for people with problematic values, to promote their own well-being is to promote their conception of a life worth living, which for Doherty is partly to live a life of sex, drugs and rock and roll. However, the phrase ‘prohibited not just for myself but for anyone’ univocally means ‘universal’. I am not denying that maybe should be the case when the loved person has a problematic identity. Loving someone does not seem compatible with endorsing their doing something self-destructive. There is a difference, however, between doing something self-destructive and having something self-destructive as a value. For Doherty, taking drugs is not merely something that he does: it is one of his values. Can Helm affirm both that the loved person’s values are subfocuses of caring and that the lover has to object to certain values of the loved person?
I do not think he can. Helm’s move to a universal standard of well-being in the context of caring about other people is in contradiction with his understanding of well-being in the context of caring about oneself. If anything can be a value for oneself, then any value that a person has needs to be a subfocus of caring about the lover. Helm has to abandon either

i. his conception of well-being as not conforming to moral or universal standards;

ii. his conception of the lover having to object to the loved person’s values when they are not conducive to a universal standard of well-being; or

iii. his conception of caring about persons as having the loved person’s values as subfocus.

I believe that Helm should abandon the latter two, given that his answer on problematic identities stems from his view’s failure to capture the role that the loved person’s values play in love. This is what Doherty’s actual partner, Jade, says about him in the second article I quoted earlier:

While Doherty disappears briefly, I ask Jade what the best thing is about him. “His soul, his inner self.” And the worst? “Probably his disbelief in himself. I think if he could love himself as much as I do ...” (Hattenstone, 2019: n.p.)

This is what Doherty’s manager, who went to school with the musician, says about him:

What was Doherty like as a schoolboy? “Utterly brilliant. He got top grades in everything. He was very similar to today apart from the obvious one that we all wish was different.” Drugs? “Yes.” (Hattenstone, 2019: n.p.)

Doherty’s value for drugs is not constitutive of his partner’s and his manager’s love for him. Jade loves Doherty despite his lack of self-respect; the manager judges him as brilliant except for his drug addiction. Universal standards of well-being matter here. It would be odd to claim that when the people we love value universally damaging
things as part of a life worth living, our love is constituted by our valuing these things. Erich Hatala Matthes expresses this worry:

> If some aspect of a person’s identity is bad for her, how can you regard it as a reason for loving her and still claim to be fully concerned for her well-being? Putting aside matters of identity for the moment, consider the following: would it be fully consistent with concern for a friend’s well-being to regard his chronic pain as a reason for loving him? Or his weakness of will? Or any other property that is bad for him? “I love that guy! He’s kind, funny, and completely lacking in self-esteem!” Those might be good reasons to love a comedian, but the last is a strange and callous reason to love a friend. (Matthes, 2016: 245)

If the values expressive of a person’s conception of life worth living are in conflict with that person’s objective well-being, loving that person cannot consist in caring about those damaging values. As Matthes points out, in these cases either we love a person regardless of her values, or we love her in spite of her values:

> To love someone in spite of an aspect of her practical identity is ... to take a property that figures centrally in a description under which she values herself and to reject it as a reason for loving her. It is not just a rejection of what she values, but a rejection of what she values about herself, of the very idea that what she values about herself is worth valuing. (Matthes, 2016: 259)

Matthes ultimately argues that although problematic values cannot ground love, we should love people for who they are even if their values are objectively bad for them (as long as they are not morally bad, like being racist) (2016: 259). My claim here is not whether we should or not, my claim is that we do sometimes love in spite of people’s values, which, as Helm rightly claims, need not be conducive to objective well-being. But we do not love them because of those problematic values, and our loving them cannot commit us to value those problematic values. So, either the loved person’s problematic values do not count towards one’s love, or instances of love in spite of are not genuine cases of love. Hence, either Helm is wrong about the role values play in love (i.e. the lover does not have to have the loved person’s values as
subfocus of the same emotions the loved person experiences) or he is denying that people with problematic values can be genuinely loved.

Then again, Helm could say in response that his claim is not that the loved person’s values ground love, but that love rationally commits the lover to a pattern of emotions towards the loved person’s values, such as shame or pity. He could say that, in virtue of Doherty’s having a problematic identity, Jade is rationally committed to being ashamed or pitiful when Doherty takes drugs even if he is not, given that he is not living as he ought. In that case, Helm’s claim \( i \) does not stand. Claims \( ii \) and \( iii \) seem to remain unaffected. As I show next, these claims do not stand either.

1.3.2. Love in spite of opposing identities

Loving in spite of may happen not only when the loved person’s values are detrimental to their well-being. It can also happen when the loved person’s values are in opposition to the lover’s. For example, think of people with different political ideologies. Journalist Lucy Mangan, who is openly pro-Labour and is married to a Tory supporter, describes their relationship as follows:

I would like, for example, to be helped with the recycling without having to listen to a diatribe about environmentalism (or “a blend of anti-human, self-praising crypto-paganism that’s gone too far”) or to watch Newsnight without explosions of rage about bias and the “bloated, Byzantine, bureaucratic nightmare that is the BBC” or have our friends round for dinner without at least half of them being served slaughtered sacred cow as the main course.

I would like to be able to give money to people who beg on the streets without my husband rolling his eyes. That one I would like very much. (Mangan, 2016: n.p.)

Mangan does see the advantages of loving someone with the opposite ideology (namely, it “banishes boredom” and it allows her to listen to different opinions to hers). But ultimately, hers is a case of loving in spite of her husband being a Tory supporter, and the other way around. It could equally be said that the husband loves
Mangan in spite of her watching the BBC, her vegetarianism and her tendency to give money to homeless people.

Which pattern of emotions could Helm argue that Mangan is rationally committed to towards the Tory party, or her husband towards the Labour party? Being a Tory or being pro-Labour cannot count as universally problematic values. Here, Helm cannot argue that they should object to their respective values. Taking the idea to the extreme, within Helm’s picture it could be argued that, for Mangan, her husband is like Doherty, due to his identity and values not corresponding with the conception of a life worth living he ought to have. Maybe Mangan could judge that her husband’s identity impacts his objective well-being negatively. However, Helm is very clear here: judging whether a value impacts the loved person’s well-being wholly in virtue of one’s own values “would be inadequately sensitive to his identity as this person and, in particular, to the place his autonomy ought to have in defining his identity as such” (Helm, 2010a: 159). Even when taken to the extreme, Mangan cannot judge her husband’s well-being in virtue of her own values. In cases of loving in spite of opposing identities, the lovers cannot object to each other’s values. Claim ii does not hold.

In order for claim iii) to hold, the only option for Helm here is to say that Mangan is committed to being joyful towards the Tory party as subfocus when her husband is joyous about a Tory victory. This is initially more plausible in opposite identities than in the case of problematic identities. It is a less outlandish claim than claiming that Doherty’s partner is rationally committed to being happy when Doherty happily takes drugs. However, this is revealing of Helm’s complete oversight of the lover’s conception of a life worth living, and how caring for another person relates to such conception.

In Helm’s account, in virtue of loving someone with values that are in conflict with her own, Mangan is committed to a pattern of emotions which goes against her own conception of a life worth living. Let us remember that, for Helm, caring about oneself is to be rationally committed to a pattern of emotions towards one’s own identity. This pattern of emotions plausibly includes sadness, not joy, about Tory victories. Being joyful about a Tory victory is a failure of her caring about her own identity, given that it is in opposition to Mangan’s value as pro-Labour. Mangan is already committed to a pattern of emotions towards the Tory party due to her own
values, so loving her husband cannot commit her to an opposing pattern of emotions towards the Tory party. That is taking the requirements of love too far and is ultimately threatening to the lover’s autonomy. Claim iii does not hold either—at least not in the case of opposing identities. As I show next, the claim does not hold either even when they are not problematic or opposing identities involved.6

1.3.3. Not caring about neutral values

In §1.1, I argued that Helm’s view is better suited than Frankfurt’s to explain what being in love is. However, I have shown that love in spite of presents problems for Helm’s view. I will show now that neutral values (i.e. non-problematic and non-opposing) present problems for the caring view in general.

The couples I discussed in §1.1 and §1.2 fit well with Helm’s concern–based identification. Manchester United is one of Aby’s values, and Adam cares about Manchester United in virtue of loving Aby—not because Manchester United is one of his values. Philosophy is one of Marta’s values, and Janis cares about philosophy in virtue of loving Marta—not because philosophy is one of his values. However, these couples need not be the paradigm from which we can extract universal claims about what being in love is.

It is all too common, in the discussion of love, to do conceptual analysis which only explains love at its best expression. Edward Harcourt takes note of that tendency:

> Philosophical accounts of love are often idealistic: if it doesn’t live up to some ideal standard, it isn’t love. Of course we can use the word ‘love’ this way, like the person who storms out of a gallery complaining that what they have seen ‘isn’t art’. But I take it that their complaint is not that what they’ve seen doesn’t fit a definition of art, but that it is bad art. Similarly, the line which all philosophers who write about love seem fated to quote is not Shakespeare’s shot at a definition of love, but his way of saying that love that alters when it alteration finds is an

6 It could be argued that Helm accommodates love in the case of opposing identities in his discussion about conflict between values (2010a: 122-12). If a person who is both a father and a professor and misses a class because he is taking his daughter to the hospital, he should not feel ashamed that he failed to uphold his valuing being a professor (2010a: 123). However, here Helm is discussing conflicts between values a person has already: being a Tory supporter is not one of Mangan’s values so Helm’s discussion on conflict between values does not apply here.
example of bad love: ‘love’ in that context – the idealizing usage – is the name of an excellence. But though love can be the making of lives, it can also be their unmaking. (Harcourt, 2016: 39).

Here, I call this idealised type of love ‘harmonious love’. In harmonious love, both the lover and the loved person have non-problematic identities, promote each other’s well-being and care about the other person’s interests. The almost exclusive use of harmonious love in the conceptual analysis of love means that the resulting concept of romantic love does not include people with problematic identities or opposing values. Ultimately, this idealising approach does not accommodate many everyday experiences of love.

In his bagpipe example, Helm is extracting a universal claim about love from an example of love that cannot be universalized. Helm feels proud about his wife’s bagpiping prize; he then claims that the lover is rationally committed to feeling proud. But would it really be irrational to not feel proud about one’s wife bagpiping prize? Think of the husband who is tired of having to go to bagpiping conventions every month and wishes his wife would, at least, swap to a less annoying musical instrument. I mentioned earlier that any philosopher with a non-philosopher partner would know that her partner only cares about philosophy in virtue of it being one of his values. This is probably the case in most non-philosopher partners who do care about philosophy. But what about those who do not care about philosophy at all? Are these people irrational? More importantly, can we still say that these people are in love with their partners? I have already shown that Helm’s view does not accommodate cases of problematic and opposing values. Now, I argue that it does not accommodate other instances of love where there are no problematic or opposing values involved. I illustrate this point with two examples where one of the partners has Manchester United as one of their values. Manchester United is, in

---

7 These cases are even less fitting with Frankfurt’s view, which would entail that drugs have to become one of Jade’s interests, and the Tory party one of Mangan’s interests. Since I already accepted in §1.1.2 that Helm’s view is more suited than Frankfurt to explain what being in love is, I focus my critique on Helm only.
principle, neither a problematic value nor a value that can be in opposition with someone else’s.⁸

In a harmonious relationship like Aby and Adam’s, Adam does care about Aby’s value of Manchester United. However, this is not the case of Angela and Silvio. Manchester United is one of Angela’s values, but Silvio does not care in the slightest about Manchester United. He pays no attention to Manchester United, he is not motivated to act about Manchester United. Not only it is not one of his values that make his life worth living: it is simply not one of his interests.

Then, we have Lydia and Mema. Mema absolutely hates football. It is not against her conception of a life worth living; she just does not like it. Manchester United is one of Lydia’s values, and that is a source of unhappiness for Mema. This case is different from loving in spite of problematic identities. It is not the case, like in Doherty’s example, that Lydia’s values are not conducive to her well-being. Mema knows well that going to matches helps Lydia unleash tension and gives her a sense of belonging: it increases her well-being. It is not a case of loving in spite of opposing identities either. Manchester United is not opposed to any of Mema’s values. Unlike Mangan and the Tory party, if Mema were to be happy about Manchester United winning, that would not be in contradiction with any of her values. It is simply the case that Mema despises football.

Silvio and Mema do not experience their respective partners’ emotions towards Manchester United in virtue of caring for their respective partners. Here, Helm has two options. He either has to say that Silvio and Mema do not love their respective partners, or he has to surrender his definition of love as caring for the other person’s identity with her values as subfocus. As I argue next, it would be hard for Helm to make the former claim—that Silvio and Mema do not love their partners.

When Angela is happy about a Manchester United victory, Silvio is neither happy nor sad about Manchester United. However, Silvio values highly that Angela is so

---
⁸ There could be instances where valuing Manchester United is problematic, if the loved person is a radical hooligan and this value is connected to violent or anti-social behaviour. In that case, being a Manchester United fan would be analogous to Doherty’s valuing drugs with respect to their well-being. It could even be an opposing value for a hardcore Manchester United fan who is in love with a hardcore Liverpool fan, given the enmity between both clubs. This would then be an analogous case to the Labour and Tory partners. These cases should be understood then as instances of love in spite of. The cases I discuss next should be understood as non-problematic and non-opposing.
passionate and committed to something that she cares about. So, when Angela is happy about a Manchester United victory, Silvio is happy. He is happy about Angela, not about Manchester United. There is no need for Manchester United to be the subfocus of Silvio’s emotions, but nevertheless, it seems like this being happy for her is a reason to believe Silvio loves Angela.

In the case of Lydia and Mema, I can anticipate Helm saying that even if they are in love, theirs would be a defective type of love. However, I think that would overlook that love is, quite often, a source of conflict. Basing a theory of love on harmonious relationships results in a failure to accommodate that people can love each other and still be in profound disagreement with each other, and that it is precisely out of love that they navigate those disagreements.

Amélie Rorty is one of the few theorists who gives conflict a place in love. We love our partners, Rorty says, but we also love our careers, our hobbies, our friends. Sometimes, these are in conflict with each other, and love “involves maintaining the dynamic equilibrium and harmony among the competing demands of multiple loves” (Rorty, 2016: 349). Lydia and Mema must negotiate in order to achieve such equilibrium: they may reach an agreement that Lydia will go on alternate weeks and go hiking with Mema every other Sunday. Lydia does not care about hiking; she would rather go to watch football every week. Mema does not care about Manchester United; she would rather go hiking every week. But, out of their love for each other, they negotiate their conflicts. theory of being in love needs to accommodate negotiation and conflict as a feature of any love, not as a feature of deficient love. Otherwise, only harmonious love, where all of the lovers’ interests overlap, would count as love. That would be too stringent, and possibly theoretically superfluous – there would not be many relationships which fulfil the conditions of the theory, so the theory would not be explaining a very interesting phenomenon.

These are examples of people who are plausibly in love with each other, but the lover is not committed to a pattern of emotions towards the loved person’s values. If being in love is possible without a rational commitment to a pattern of emotions towards the loved person’s values – which I believe it is in light of the examples I have given above – then the rational commitment of emotions towards the loved person’s values is not a necessary feature of being in love.
On that basis, I believe that the caring approach to love is misguided, and propose a change of direction.

1.4. Making love about the lover

Despite not being able to accommodate non-harmonious love, the caring view does get some things right. As much as conflict may be constitutive of love, there must be some degree of promotion of the loved person’s interests for a relationship to merit the label of loving. I am not refusing that loving someone entails or maybe even commits one to a pattern of emotions towards certain things in the exact sense described by Helm. It would not be very loving to not feel sadness when one’s partner is sad due to the death of a friend. It would not be loving either to be indifferent towards all of one’s partners successes, or to not be happy for them. Going back to the very first example, if the only reason Manuela is happy about Casey’s promotion is that she will be able to go on holiday to the Maldives, we would be right to doubt whether she really loves him – at the very least we would not count this episode as counting towards accepting that Manuela is in love with Casey. The problem for Helm is that he defines love in virtue of this commitment, making it rationally required.

Also, when we are in love, we tend to be attentive and active towards the people we love: we notice things related to them (things about them, or things that they value); we act to promote these things. Silvio may not care about Manchester United, but nevertheless will make sure that he does not schedule a family dinner for the night of the Champions League final: he will notice that is an important day for Angela and act in consequence. The notion of love involving attention and action on behalf of the loved person is also a right assumption of the caring approach. But then again, defining love in virtue of this commitment fails to explain everyday experiences of the phenomenon.

Furthermore, behind the caring approach (as well as union theories) there is a right intuition about how people’s identities are formed. I briefly mentioned in §1.2 that one of Helm’s criticisms of traditional views of love is that they have an individualist conception of persons, which assumes that to delegate responsibility for one’s own identity is inauthentic (2010a: 11). Helm argues that this conception of persons ignores “the way particular relationships can significantly enhance our lives, our
activity, and our autonomy by dissolving the social barriers that separate people from each other” (Helm, 2010a: 14). Helm believes that the lover and the loved person eventually end up sharing a single evaluative perspective. The rejection of his view on caring and values invalidates this conclusion, which is probably excessive. Nevertheless, it highlights another crucial feature of love that the caring approach gets right: being in love entails a change in the lover.

An analysis of being in love should preserve these intuitions about love without making them necessary conditions, or defining love through them. I believe that this feature of the caring view rests on the fixation of traditional views of love with defining love in terms of caring for the loved person for the loved person’s sake. This has translated into a generalized acceptance of disinterestedness as a necessary feature of love, and as an establishment of the loved person as the main focus in conceptual analysis. As seen in the introduction, the disinterestedness condition even leads Frankfurt and Velleman to reject romantic love as a pure form of love. However, the disinterestedness condition is increasingly been put into question. Neil Delaney (1996), Edward Harcourt (2011), Gary Foster (2008), Sara Protasi (2016) and Monique Wonderly (2017) have all defended that there is at least a degree of self-interest in love. Chesire Calhoun rightly points out that one type of reasons for love are what she calls ‘reasons-for-me’:

[R]easons-for-me… is a large and diverse category of reasons. It includes the symbolic value an activity has for an individual, its biographical significance given his past, its suiting his personality and natural talents, the fact that he made a promise or that people important to him would

---

9 Monique Wonderly and Agnieszka Jaworska’s Caring and love (forthcoming), unpublished at the time of writing, note the same flaws and strengths of the caring view that I have developed in this chapter. My account is broadly compatible with Wonderly and Jaworska’s and certainly shares many points in common with it. For example, according to Wonderly and Jaworska, a lover’s sense of herself as an agent leading a meaningful life is directly impacted by the faring of the loved person, without the need to share the loved person’s interests. This is a foundation of the arguments in this chapter and chapter 5. Nonetheless, amongst other significant differences, my account is most notably distinct from Wonderly and Jaworska’s in that we ultimately have different aims. While Wonderly and Jaworska are interested in establishing the necessary and sufficient features to distinguish between loving and caring, I am interested in defining what being in love is. In that sense, our accounts diverge once we lay out our common criticism to the caring view -which we also cash out in different terms: me in terms of loving in spite of and them in terms of instances of caring that are not instances of loving. I am grateful to Monique Wonderly for making their manuscript available to me in advance.
be disappointed in him if he didn’t select this activity as an end, and that he enjoys it. (Calhoun, 2016: 277)

An analysis of what being in love is needs to accommodate this element of self-interest in romantic love. Hence, such analysis has to take into account the lover as well; after all, the lover is the person who is in love. The lover is the person whose interests, attention and action change in virtue of being in love, as the caring approach acknowledges. I come back to the issue of self-interest in §5.1.2. For the moment, I think my argument about the caring view warrants my change of focus from the loved person to the lover.

**Conclusion**

In this chapter, I have described Frankfurt’s and Helm’s views of love as caring. For Frankfurt, to care about someone entails an incorporation of the loved person’s interests into one’s own. For Helm, to care about someone entails a rational commitment to a pattern of emotions towards the loved person, with the loved person’s values as subfocus. I have shown that both these views are mistaken, given that they cannot accommodate cases of non-harmonious love, where either there are problematic or opposing identities, or the lover does not care about the loved person’s neutral values. Ultimately, I have suggested that in order to keep the right intuitions of the caring approach, we should change the focus. Instead of analysing love by trying to discern what it is to care about a person for her own sake, I propose that we investigate the change that being in love entails for the lover.
Chapter 2

Narrative theory

In order to explain what being in love is, I focus on the lover’s self-concept. My final aim in the thesis is to offer support for the following claim:

Being in love is a defining trait of the lover’s narrative self-concept, which articulates the lover’s understanding of herself and the world.

Situating the discussion at the level of the self-concept allows me to explain what being in love is without the need to make claims on autonomy, morality or personal continuity. I focus on a specific approach to the self-concept: narrative theory. Because of the need to distinguish the self-concept from personal identity and due to the fact that narrative theory can be understood in multiple ways, this chapter is aimed at situating my general approach and setting up a plausible view on the narrative self-concept—which I develop in chapter 3.

In §2.1, I make two preliminary distinctions. Firstly, I justify my departure from the philosophy of emotions. Secondly, I justify my departure from the metaphysics of personal identity. I define the self-concept and explain it is the most adequate framework for the project.

In §2.2, I map out narrative theory. I distinguish two clusters of views within narrative theory: Strong and Minimal Narrativism. Strong Narrativism requires a coherent narrative of one’s life for personhood and/or morality. Minimal Narrativism, for its part, considers narrative thinking as a tool for understanding oneself and the world.

In §2.3 I present the most commonly held objections to narrative theory, synthesised in Galen Strawson’s Against Narrativity (2004). Some of Strawson’s anti-narrativist challenges only hold against Strong Narrativism, but they are useful to set limits on any plausible narrative theory.

Finally, in §2.4 I explain how the anti-narrativist challenge fares against Minimal Narrativism. Daniel Hutto’s (2007, 2016) and Peter Goldie’s (2012) views are open to Strawson’s objection from triviality: if form-finding (i.e. creating stories about
oneself) is not necessary for self-understanding, then narrative theory is trivial. Developing a non-trivial account of the narrative self-concept will be the aim of chapter 3.

2.1. Preliminary distinctions

2.1.1. Departing from the philosophy of emotion

Once I have taken the step of putting the focus on the lover, one plausible route would be to define being in love in terms of the lover’s mental states. Following this route, this thesis would construe being in love as an attitude, and more specifically, as an emotion. This is not the route I will follow.

There have been several attempts to define love as an emotion. Emotion theories have difficulty explaining certain peculiarities of romantic love that make them different from more straightforward or basic emotions, like fear. Arina Pismenny and Jesse Prinz summarise some of romantic love’s peculiarities:

First, love can be felt. People report being overwhelmed by the feeling of love. But does this mean that once in love, one is constantly experiencing the feeling of love? It seems not, for one may be in love for years, and not feel love 24/7. This suggests that love is not simply a feeling but instead also exists in a dispositional form. Second, the feeling of love may fall anywhere on the spectrum of pleasantness and unpleasantness. Love can be extremely elating, taking the lover to the highest heights of her happiness, and it can also bring her down into the deepest void of despair (Pismenny and Prinz, 2017: n.p.).

To my knowledge, no one has seriously attempted to define romantic love as a bodily feeling. That view, as Karen Jones puts it, is “a nonstarter” (2008: 276), given that there is not one thing that love feels like: Lovers can experience a whole constellation of feelings, from joy, to despair, to jealousy; nor must the joy that lovers feel always feel the same (Jones, 2008: 276).

Most theorists of emotion have therefore followed the route of defining romantic love as a combination of emotional attitudes, which has received different labels: a reactive attitude (Abramson and Leite, 2011), a disposition (Naar, 2013) or a
syndrome (De Sousa, 2015; Pismenny and Prinz, 2017). I am, however, not concerned with the metaphysics of emotional phenomena. The main reason to depart from the philosophy of emotion is that, ultimately, I am not concerned with the same questions philosophers of emotions are concerned. Hichem Naar describes well the focus of a significant number of philosophers of emotion regarding romantic love (including his):

The question is rather whether there are considerations which are in a certain sense internal to love, whether there are considerations which make love an appropriate, merited, or fitting response to its object, a similar way there are considerations which make paradigmatic emotions—such as fear—appropriate, merited, or fitting responses to their object. The question, then, is whether there are any considerations making love fitting a similar way the fact that a given situation poses a threat to one makes it fitting for one to be afraid of it. For any given individual, is there a fact of the matter whether or not it is fitting to love them in a similar way there is a fact of the matter whether or not it is fitting to fear them? (Naar, 2019: n.p.)

So, even though it is widely recognized that love is more complex than fear, there is still a tendency to ask the same questions we ask about fear; namely, whether love can be rationally justified and if so, on what grounds. As I said, in this thesis I depart from the reasons of love approach in order to answer the question of what being in love is. Although philosophers of emotion ask what being in love is—or, at least, what love is—they do it in terms of reasons. Hence, my question is only answerable if we refuse to reduce it to a question regarding emotions.

Obviously, I do not deny that being in love consists, at least in part, in being in some emotional state or states. My strategy is to assume that romantic love is at least partly affective, without taking a stance on whether romantic love can be defined through a specific emotion label (a feeling, a disposition, a reactive attitude, a syndrome) or how can it be rationally justified. Some of the theories I draw from do discuss romantic love in terms of emotions. My aim is to build upon the claims they make without committing to their position in the context of the emotion debate. This is a valid strategy because my thesis is concerned with the lover’s identity—not the lover’s attitudes, although these will be irremediably related.
In that sense, my project starts from Helm’s approach, notwithstanding the objections I have made in chapter 1. Helm defines love as a rational pattern of emotions, but mental states are not the foundation of his theory; personal identity is. Let us remember that loving oneself is to be rationally committed to a pattern of emotions towards one’s identity, understood as a general conception of a life worth living, which is interdependent with one’s values. An implicit consequence of Helm’s view on identity is that, in virtue of being in love, the lover acquires a new value which is related to her own conception of a life worth living. Some of the values Helm mentions are instantiated by roles (being a professor) and relationships (being a father). Being in love with someone is potentially realised through being someone’s partner, and that can be interpreted as a role or a relationship. Being someone’s partner can then be a value for the lover: it can become part of her conception of a life worth her living; part of her identity.¹⁰ That will be the approach I follow, which requires a second clarification.

2.1.2. Departing from the metaphysics of personal identity

Helm’s view on love follows from his specific view on personal identity. For Helm, a person’s identity is defined by her general conception of a life worth living and the values that relate to that conception. Persons have the capacity to determine and evaluate their own identities, which are then continuously subject to change. One does not determine one’s identity in isolation, given that people who are close to one—such as friends or romantic partners—contribute to it. In Helm’s view, this happens through the acquired patterns of emotions towards the loved person’s values. But I have rejected that approach, so I will try to offer an alternative explanation. Before doing that, there is a question about personal identity that I shall leave aside.

In metaphysics, the term ‘identity’ commonly refers to numerical sameness: is this one thing the same as this other thing? It makes certain sense then that personal identity has been approached from a numerical perspective: is this one person the same as this one person? In the case of persons, current numerical sameness is relatively unproblematic: ‘I am identical to myself now’ is prima facie uncontroversial.

¹⁰ Helm recognizes this change, given that ultimately his view is that the lover and the loved person share a single evaluative perspective. But he never acknowledges that the loved person becomes a value for the lover, probably because that claim sounds too close to Frankfurt’s incorporation of the loved person’s well-being to one’s own. Those suspicions do not affect my view, since I have already rejected the caring approach that entails such incorporation.
That means that the focus of the discussion has often been on numerical identity of persons over time. I explain the approach from a widely used example put forward by Derek Parfit:

*The Nineteenth Century Russian.* In several years, a young Russian will inherit vast estates. Because he has socialist ideals, he intends, now, to give the land to the peasants. But he knows that in time his ideals may fade. To guard against this possibility, he does two things. He first signs a legal document, which will automatically give away the land, and which can be revoked only with his wife's consent. He then says to his wife, ‘Promise me that, if I ever change my mind, and ask you to revoke this document, you will not consent.’ He adds, ‘I regard my ideals as essential to me. If I lose these ideals, I want you to think that I cease to exist. I want you to regard your husband then, not as me, the man who asks you for this promise, but only as his corrupted later self. Promise me that you would not do what he asks.’ (Parfit, 1984: 328)

Parfit’s aim is to determine whether the young idealist really ‘ceases to exist’ and is substituted by a new person, the accommodated old nobleman. The view is based on Locke’s claim that what warrants the survival of a person over time is the psychological continuity of consciousness (1694 [1998]). The idea is that personal identity depends on the existence of psychological connectedness between the mental states of two temporally distant iterations of a person. A disruption in psychological connectedness changes who you are, and ‘who you are’ is understood as a collection of mental states.

Marya Schechtman renamed the question of continuity as the *reidentification question.* As she describes it, reidentification defines the identity relation that an object—in this case, a person—bears with itself over time (Schechtman, 1996: 75). Schechtman coined the term to contrast it with another question—in her view, one that was being neglected in the debate. Schechtman calls it the *characterisation question*: the definition of “the set of characteristics that make a person who she is” (1996: 75–76). For Schechtman, there are “actions, experiences, beliefs, values, desires, character traits” (1996: 73) that make a person who she is. In the case of Parfit’s Russian, the characteristic ‘being a socialist’ makes him who he is, so losing that characteristic means he stop being who he is—he becomes a new person. Broadly understood, the
change of characteristics changes who you are, and ‘who you are’ is understood as a collection of actions, experiences, beliefs, values, desires and character traits.

Reidentification and characterisation are two separate questions about personal identity, and are relevant in different contexts. Schechtman gives an example of a situation where reidentification is irrelevant: buying a sports car. A person may be, in fact, a different entity if Parfit is right about psychological connectedness, but the car loan still needs to be paid by that person (Schechtman, 2007: 170). Equally, the young idealist and the accommodated old nobleman may not be the same metaphysical entity. But if the nobleman had an unpaid debt from his idealist years, his claim that he is a different metaphysical entity would not go very far in court. What these examples reveal is that the ontological status of a person over time has no bearing on certain features of that person’s life—in these cases, on legal responsibility. Romantic love is another one of those features.

My view is formulated within the characterisation question. The aim of this thesis is to show that romantic identification is the process of acquiring the characteristic of being in love. The metaphysical status of persons after change does not reveal anything interesting about this feature of a person’s life. Just to make it explicitly clear: I will ignore the reidentification question in this thesis, even in cases where the authors I discuss have made claims on that question.

2.1.3. The self-concept

With the reidentification question set aside, I can focus on characterisation, which is the most relevant question for my purposes. For Schechtman, there are certain “actions, experiences, beliefs, values, desires, character traits” (1996: 73) that make a person who she is. This definition is very abstract and seems to overgeneralise. I am now performing the action of writing this chapter: does that really ‘make me who I am’? It does not seem so. I can have a belief that the Earth is round; I can value money as an instrumental way to survive in society; I can desire a green pair of shoes for my birthday. These are not interesting answers to the question of who I am. Even character traits are not immediately illuminating for this questions: I can be sociable without that trait really defining who I am. It follows then that only some actions, experiences, beliefs, values, desires and character traits are characteristics, and others are not.
Schechtman calls characteristics that make a person who they are “robust inclinations” (2004: 416). She establishes three necessary features of robust inclinations: (i) they should be stable, coherent, and powerful; (ii) they should not originate from a physical or psychological pathology; and (iii) they cannot be merely desires for a state of affairs, “but also for a way of being or a type of life” (Schechtman, 2004: 416). In this sense, Schechtman’s characteristics mirror Korsgaard’s practical identities – descriptions “under which you value yourself and find your life worth living and your actions worth undertaking” (Korsgaard, 1996: 101). And they are also similar to Helm’s values. In fact, Helm’s values only differ from Korsgaardian practical identities in one respect: Korsgaard believes that a person’s identity is constituted by practical identities, while Helm believes that identity is both prior to and interdependent with values. With the concept of a life worth living in hand, it is easier to see what characteristics or expressions of characteristics are.

Is the action of writing this chapter part of what makes me who I am? I am a philosopher, and being a philosopher is constitutive of a life worth living for me. Hence, writing this chapter is the expression of one of my characteristics. It does not by itself make me who I am, but it is an expression of who I am. If I had commissioned this chapter to a ghostwriter who only does it to pay the rent, the same action performed by him would not be an expression of a characteristic in his case. Is the belief that the Earth is round part of what makes me who I am? No. But it could be, if what made my life worth living was that I am a hacktivist dedicated to debunking online conspiracy theories (like that of Flat-Earthers). Is the valuing of money as an instrumental way to survive in society part of what makes me who I am? Not really. But it can be a characteristic for someone who dedicates their life to campaigning against crypto-currencies. Finally, does the desire of getting a pair of green shoes for my birthday make me who I am? Although I do want a pair of green shoes, it is not. But imagine an Instagram influencer whose trademark is green shoes: it could be an expression of her characteristic ‘being an Instagram influencer’. A parallel can be drawn with respect to character traits: being sociable may be just another part of my personality, but if I am the type of person whose life is heavily focused on social interactions, it can be one of my characteristics.
In that sense, corresponding with what I have said in the previous chapter, I am going to follow Helm’s separation of characteristics from universal or moral standards. The above examples can be re-formulated with characteristics which do not conform to these standards: being a mafioso, believing that men are superior to women, valuing violence, or desiring for the world to burn can all be people’s characteristics or expressions of them. If I am to claim that being in love is a characteristic, and it is possible to be in love with disvaluable people, I must have a theory that is compatible with the view on characteristics as not conforming to universal or moral standards. However, this picture of characteristics would not be widely endorsed. I said in chapter 1 that Frankfurt and Korsgaard would reject Helm’s view on well-being. I also showed that Helm’s theory of well-being is put into question by his own account of caring for persons. Schechtman, would also plausibly object to Helm’s theory, because she sees personhood as inseparable from morality (2007: 164). These authors discuss characteristics in terms of personhood or selfhood, which in turn are inseparable from agency and autonomy. It is hard—if even possible—to separate agency and autonomy from universal and moral standards.

To defend my view, I will fall back on an idea which can be plausibly discussed independently from agency and autonomy: the self-concept. Psychologist Roy F. Baumeister defines self-concept in the following terms:

The self-concept is the individual’s beliefs about himself or herself, including the person’s attributes and who and what the self is. The self-concept includes many things that might not be part of one’s identity. For example, a person’s self-concept might include many personality attributes, such as being friendly or talkative. (Baumeister, 2005: 247)

When Baumeister says that some components of a person’s self-concept are not part of her identity, what he means is that they are not related to her metaphysical status as a specific entity (in this case, a specific person). Looking at his definition of identity sheds light on that claim:

Your sense of identity refers to your knowledge of who you are. Identity always answers the question, “Who are you?” Self-concept, in contrast, may contain answers to other questions like “What kind of person are
Identity means being the same person you were yesterday or ten years ago; it also means being different from someone else. (Baumeister, 2005: 248)

In distinguishing between identity and self-concept, Baumeister is saying is that there are features of our self-concept that have no bearing on metaphysical identity. I do not need to subscribe to Baumeister’s wider account on the self-concept, identity and the self. All I need to point out is that I situate my thesis at the level of the self-concept, and that the self-concept can be discussed separately from personal identity.

I follow Helm in his claim of a person’s conception of a life worth living being both a matter of “autonomous invention and rational discovery” (Helm, 2010a: 131). By autonomous invention, Helm means that a person’s autonomy is precisely her capacity to determine her own values. However, one can be mistaken as to one’s values, and hence self-understanding is also a matter of discovery: I may discover one day, for example, that being a philosopher is not actually one of my values. The possibility of being wrong about one’s values leads to a further clarification. Self-understanding is not reducible to self-knowledge. Firstly, one can have false beliefs about oneself; that is, one can have defective self-understanding. This is not a problem for my view, since I am not trying to establish the truth conditions for a person’s belief that she is in love. That is an interesting question which has largely been ignored by analytic philosophers of love, mostly focused on defining necessary reasons for love, not explaining how someone can access those reasons. I will, however, continue the trend, and remain vague as to which are the truthmakers for the belief that one is in love. I assume that it is possible to believe one is in love and, at some point, realise that one was wrong all along; this is not infrequent.

Secondly, and relatedly, one can ignore certain things about oneself, or have very unstructured, changing beliefs about oneself. Again, this is very common in the case of being in love. One may not know one is in love, or be unsure about it, or change one’s mind about it during a given period of time. The fact that we can have defective self-understanding does not open objections to my view: on the contrary, it offers support for it versus views where it seems that it is always clear cut whether we are or not in love.
Once these distinctions have been made, I can move on to the specific framework I use to define the self-concept: narrative theory.

2.2. Different formulations of the narrative claim

2.2.1. The basics of narrative

The word ‘narrative’ may refer to different concepts. The most basic definition of the term is that a narrative is a story, understood as a representation of events. Peter Lamarque points out that “at least two events must be depicted in a narrative and there must be some more or less loose, albeit non–logical, relation between the events” (2004: 394). As Lamarque explains, the term ‘narrative’ can also refer to the text of a written story or to the act of narration or telling the story, with the story itself being a narrative product. Moreover, Peter Goldie speaks of a narrative structure, which refers to the specific form that the narrative product takes (for Goldie, a form with “coherence, meaningfulness, and evaluative and emotional import” (2012: 3)) through the narrative process.

To simplify things, I am going to use the term ‘narrative’ to refer to the narrative product: a relation of events which acquires its form through a process of creation where these events are connected while being given a narrative structure. Peter Goldie gives more details about this:

> A narrative or story is something that can be told or narrated, or just thought through in narrative thinking. It is more than just a bare annal or chronicle or list of a sequence of events, but a representation of those events which is shaped, organized, and coloured, presenting those events, and the people involved in them, from a certain perspective or perspectives, and thereby giving narrative structure—coherence, meaningfulness, and evaluative and emotional import—to what is related. (Goldie, 2012: 2)

Broadly then, the narrative process consists in the creation of a narrative product: the selection of events which are interpreted and portrayed with a specific shape. I go back to the properties of narratives in chapter 3, but this serves for the moment as a working definition.
The main idea behind narrative theories of identity is that narratives are a route for self-understanding. Lynne Rudder Baker summarises what I call ‘the basic narrative claim’ as follows:

We persons have a keen (if not universal) desire to make sense of ourselves. One common way to do this is to construct stories about our lives and the connections between what we have done and what has happened to us. These self-narratives are autobiographies that we – perhaps without any intention on our part– use to give meaning and coherence to the events of our lives. (Rudder Baker, 2016: 7)

The claim that we create stories about our lives has different interpretations. Depending on their stance, narrative theories can be grouped under the labels of Strong Narrativism and Minimal Narrativism.11 The first group operates at the level of selfhood or personal identity; the second group operates closer to the level of self-concept. What they all have in common is that they all endorse the idea that self-understanding is historical: we understand events in our life and thoughts that we have in light of each other. Furthermore, all narrativists implicitly or explicitly endorse the idea that identity/the self/the self-concept is formed relationally, in interaction with other people. These two pillars of narrative theory situate it in equal terms with Helm’s theory of love. Helm rightly recognizes that love is historical, since he establishes a pattern of emotions as a necessary condition for love. A pattern of emotions can only be understood as something that happens across time; so for Helm love is historical. This makes sense: poetic licences aside, it seems odd to say that one is in love ‘for a moment’. Furthermore, Helm rejects the individualist conception of autonomy and calls for a recognition of how identity is formed relationally, particularly through close relationships like friendship or romantic love. I will properly develop the historical and relational features of love in chapters 4. For now, it suffices to say that narrative theory seems prima facie compatible with

---

11 There are other taxonomies of narrative views (Schechtman, 2007), and some of the authors I discuss below place themselves in a different cluster than I do here. This is more frequent among Strong Narrativists: Schechtman considers her view a “middle-range” account (2007: 163), while Anthony Rudd, whose account Schechtman considers a strong one, argues that “[w]e are… all narrative in the middle-range… sense” (2012: 207). Here I adopt Daniel Hutto’s divide (2016: 26-29), given that it is done from the Minimal Narrativist approach that I myself will follow. Also, I drop Hutto’s term ‘Weak Narrativism’ and use ‘Minimal Narrativism’ instead because my aim is to give the minimal requirements for a view to be considered narrative –while Hutto is more worried in distinguishing his view and Goldie’s from Strong Narrativism.
those two features of love also recognized by Helm – and for that reason it is the framework for my view on being in love based on the self-concept.

As I have just mentioned, however, there are a myriad of formulations of the basic narrative claim with very different metaphysical and ontological requirements, which tend to be grouped as one. Some of these views are too demanding and cannot plausibly be the framework for my view, so I need to offer a plausible narrative theory before I can offer an account of the narrative self-concept. In the remainder of the section I describe the main features of Strong Narrativism and Minimal Narrativism. This section is not intended to be a detailed discussion of the accounts portrayed, since it is a further step to properly situate my view.

2.2.2. **Strong Narrativism**

There is a cluster of views which use the basic narrative claim to define necessary conditions for personhood and/or moral agency. The views in this cluster – which I label Strong Narrativism – share the idea that one’s actions and beliefs must form a coherent narrative overarching one’s whole life – or, at least, extended pieces of one’s life. They also concur with the claim that we are the authors of such narratives and, by extension, of ourselves: by giving narrative form to the events in our lives, we are in fact shaping or constituting who we are. In this group, there are authors who see the narrative claim as an ethical prescription, i.e. narrative coherence is morally required. Others are not so committed but subscribe to a psychological or metaphysical prescription, i.e. narrative coherence is a requisite for personhood and agency. I briefly describe both versions of Strong Narrativism next.

Anthony Rudd synthesizes the first group’s core claim when he says that

> [a] serious self-evaluator must understand his or her life as a narrative unity, and act so as to shape and direct its continuing narrative—to be in part the author of his or her own story. (Rudd, 2012: 202)

---

12 The use of the expression ‘constitution of ourselves’ as synonymous to ‘constitution of who we are’ is intentionally vague. Most Strong Narrativists would probably have used a more precise ‘constitution of our selves’, which treats the self as a distinguishable entity. Since I have situated my view at the level of the self-concept, I do not need to discuss this element of narrative theory - I will anyway reject Strong Narrativism as a plausible route to explain the narrative self-concept in §2.4.
Evaluation (of actions, of beliefs, of lives) is, according to Rudd, concerned with “being for the Good” (2012: 202), where ‘the Good’ is understood as that life in which one achieves “full selfhood” (2012: 127). Personal identity (‘selfhood’) is understood in an ethical sense; as a response to both the questions ‘what kind of person should you be?’ and ‘what is the best life you can live?’. The answer, says Rudd, is to be the kind of person whose life is understood ‘as a narrative unity’. Similar claims have been made by Alastair McIntyre (2007), Paul Ricoeur (1991) and Charles Taylor (1989). The three of them coincide in considering a person’s life a ‘quest for narrative’: an unfolding, self-created story which moves “towards a climax or telos” (McIntyre, 2007: 217). In Schechtman’s words, this first version of Strong Narrativism prescribes that “a person must actively and consciously undertake to understand her life as a story, with a unified theme and little or no extraneous material, if that life is to be meaningful” (2007: 160–161).

Schechtman rejects that version of Strong Narrativism: she does not believe that people have the “clear sense of the structure or direction of their lives” (2007: 161). Instead, she vows for a weaker understanding of narrative unity according to which someone needs a certain understanding of how the events in her history hang together, an understanding that is mostly implicit but which she can access locally where appropriate, if she is to be able to engage in person-specific activities on which we place great importance —that it is necessary, for instance, for autonomy, moral agency, prudential reasoning or other kinds of higher-order capacities. (Schechtman, 2007: 161)

However, Schechtman’s view is still a Strong Narrativist one. Firstly, there is the fact that she considers narrative unity a requisite for personhood —understood as the ability ‘to engage in person-specific activities’. That means that a breakdown in narrative unity, or the absence of it, results on a breakdown or absence of the necessary conditions for being a person, such as agency or autonomy.

Secondly, Schechtman believes that in order for a person to have ownership of her own experiences, she needs to experience them as embedded in her past memories and her future projections:
We experience the present in the context of a larger life–narrative. In order to have a narrative self-conception in the relevant sense, the experienced past and anticipated future must condition the character and significance of present experiences and actions. When I have a self-constituting narrative, what happens to me is not interpreted as an isolated incident, but as part of an ongoing story. (Schechtman, 2007: 162)

Schechtman requires the understanding of one’s life as a narrative unity for personhood, and the experiencing of one’s life as a narrative unity for achieving such understanding. Narrative unity is then required for self-understanding and for self-experience. Schechtman establishes these requirements in a way that distinguishes her from the first version of Strong Narrativism.

The requirement for self-understanding in Schechtman is less stringent than the first version of Strong Narrativism. As seen in her first quote above, Schechtman’s narrative theory requires merely ‘a certain understanding’ of the narrative connection between events in one’s life where the strongest version above required active and continued creation of such a connection. She also allows this understanding to be ‘implicit’ and accessible if prompted, as opposed to the stronger requirement of it being conscious.

The requirement for self-experience is also weaker. Unlike the first group of Strong Narrativists, Schechtman does not require that one experiences one’s life as a whole with unique narrative unity. She distinguishes between ‘self-narratives’ and ‘person–narratives’ to overcome what she sees as a flaw in the view. A person narrative explains people’s actions and thoughts and how they relate to past and future actions and thoughts (Schechtman, 2007: 172). Having a unified person narrative is the condition for being the same metaphysical entity over time (Schechtman, 2007: 170). It is thus the unit which measures personal continuity, understood from the reidentification question. Self-narratives are, also, explanations of an individual’s actions and thoughts, but they can be shorter than person narratives and have no bearing on personhood. There can be several self-narratives within the life of a single person (Schechtman, 2007: 176). These units are defined by someone’s characteristics which can change over time—it is the answer to the characterisation
question. So although she does establish a requirement for experiencing oneself as unfolding through time, it need not be with respect to one’s life as a whole.

2.2.3. Minimal Narrativism

For the second group of narrative accounts, the narrative claim is a descriptive claim about how people are—specifically, how people think about or understand who they are. I call this group Minimal Narrativism, which has its most prominent proponents in Peter Goldie (2012) and Daniel Hutto (2008, 2016). Both see narrativity as a widely generalized—not required—tool for understanding oneself and the world—not a condition for moral agency or personhood.

Goldie presents his view as fundamentally distinct from the proposals put forward by Strong Narrativists. From the onset, Goldie explicitly rejects that narrative unity is required for personhood or morality (2012: 2). Where Strong Narrativism focuses on unity, understood as narrative coherence in the connections between our life events, Minimal Narrativism merely points out that creating those connections is something that we do through narrativity.

Goldie explains that narratives are tools for thinking about oneself and others through what he calls ‘narrative thinking’: “memories of our past, our deliberation and plan-making about our future, infuse our understanding of ourselves, and of others, in the present” (2012: 116). Specifically, he believes that narratives explain reasons for action:

Narratives often provide explanations of why someone had a particular motive, or why someone has a particular character or personality trait, or why someone was drunk, depressed, or angry. And the explanations that we get are narrative—historical explanations: they locate the motive,

---

13 Schechtman believes that narrative unity in self-narratives is achieved through what she calls ‘empathic access’: phenomenal access and sense of mineness of past characteristics, together with moral endorsement of them memories (2001). In later writings, she drops the requirement for endorsement and argues that empathic access should be understood as “heuristics” (Schechtman, 2016: 28). She also qualifies her requirement of narrative unity for self-understanding, clarifying that the ‘condition for personhood’ requirement is not normative. I believe that her re-formulation moves her view closer to Minimal Narrativism. However, most—if not all—discussions of her view—including Strawson’s anti-narrativist challenge I discuss in §2.3—are either previous to that re-formulation or plainly ignore it. Re-categorising Schechtman’s view would require a specific focus on her view and the way it has been interpreted that is completely unrelated to the matter at hand. I am left with no option but to keep labelling her view as a Strong Narrativist one, with this caveat.
the trait, the undue influence on thinking, within a wider nexus, in a way that enables us to understand more deeply why someone did the thing that they did through appeal to aspects of their personal history or circumstances. (Goldie, 2012: 20, emphases added)

However, unlike Strong Narrativists, Goldie does not think that narratives are required for understanding. First of all, not all events in one’s life can be neatly explained. “Life is full of surprises; people can do things that even they cannot make sense of” (Goldie, 2012: 22). Goldie gives the example of Joseph Conrad’s *Lord Jim*. The title character jumps off his ship without either the character or the narrator being able to explain why he did it; it “happened somehow” (Conrad, 2002: 81 cited in Goldie, 2012: 22). There is no explanation available for this chain of events. Secondly, aiming at finding a narrative explanation for everything, he says, is potentially damaging. Goldie warns that is one of our “dangerous” fictionalizing tendencies (2012: 151) to try to find an explanation at all costs, just because the absence of an explanation leaves one with uncertainty. He especially emphasises the dangers of trying to find agency where there is none. Hence, despite Goldie’s endorsing of narrativity as a tool for understanding oneself and others, it is not a requirement—it can even be pernicious to see it as such.

Whereas Schechtman says that self-narratives may be implicit but must be accessible, Goldie offers a different approach to the implicitness of narrative thinking. What he calls ‘narrative thinking through’ can take a number of shapes, some of them not involving the creation of stories whatsoever:

I can think through a narrative in a variety of ways…: remembering, perhaps episodically or experientially, some events that happened to me or that I observed; hypothesizing what might happen in the future; thinking through a narrative propositionally; mentally rehearsing it, as one might mentally rehearse a speech; using perceptual imagination to, for example, visualize what might happen; thinking through an episode of events coolly with little or no emotional engagement; thinking through the episode vividly and highly emotionally; thinking of ‘that terrible quarter of an hour when I thought that all was lost’; thinking through how things might have turned out differently if I hadn’t made
that terrible mistake; and in lots of other ways, many of which cross-cut with each other. (Goldie, 2012: 4).

These forms of reflection, e.g. episodic remembering or visualizing, entail no explicit connection between events. The question then arises of how they can be narrative. Goldie believes that we have a ‘narrative sense of self’: “the sense that one has of oneself in narrative thinking, as having a past, a present, and a future” (2012: 118). That sense of oneself pervades narrative thinking, and gives a narrative character to iterations of thinking that cannot be called ‘a narrative’ in themselves. It is the sense that we have of ourselves as enduring through time what gives those iterations of thinking a narrative character. The narrative sense of self informs not only the way we think about ourselves, but also the way we think about other people, and about the world in general (Goldie, 2012: 119).

Goldie’s claim is then about self-experience as well. However, he firmly rejects the Strong Narrativist version of that claim: “[L]ife is not a narrative” (Goldie 2012: 162). Narrative unity, understood as experiencing one’s present embedded in one’s life as a whole, is discarded by Goldie and it should not be understood in those terms.14 His claim on self-experience is a very minimal claim about the phenomenology of perception. He summarises it by drawing from Peter Strawson:

It seems…not too much to say that the actual occurrent perception of an enduring object as an object of a certain kind, or as a particular object of that kind, is, as it were, soaked with or animated by, or infused with…the thought of other past or possible perceptions of the same object….Non–actual perceptions are in a sense represented in, alive in, the present perception. (Strawson 1974: 53, quoted in Goldie, 2012: 119)

In other words, the narrative sense of self is merely a sense that the past is alive in the present, not a requirement to experience oneself as an enduring being over time.

14 Goldie insists that having a narrative sense of self is not the same as having ‘a narrative self’. He sees the idea of ‘narrative self’ as tied not only to Strong Narrativism, but to the reidentification question, which he thinks is not a relevant framework for narrative theory (he focuses specifically on Schechtman’s requirement of empathic access for personal continuity in 2012: 128-141). Since I have already set reidentification aside, I do not need to engage with this distinction. But it should be noted that Goldie’s narrative sense of self is specifically posited in opposition to the claim that narrative plays a role in personal continuity.
Hutto also departs from Strong Narrativism. Hutto’s formulation of the narrative claim is very similar to Goldie’s:

[N]arratives play a central role in making it possible to evaluate our reasons for acting by reflecting upon our impulses, dispositions and attitudes. Narrative capacities make this possible by providing a way for us to think about the particularities of our characters and situations—what has gone on in our pasts and what might occur in our possible futures. (Hutto, 2016: 25)

Hutto focuses on what he calls “the narrative basis of our folk psychological competence” (2008:22). He claims that ‘story-telling practices’ are “the normal route for acquiring the competence needed for making sense of ourselves and others in terms of reasons” (Hutto, 2016: 39). Like Goldie, Hutto sees narrativity as a tool for understanding oneself and others. Also like Goldie, Hutto does not claim that narrativity is a required tool for understanding, or that narrative self-understanding is universal.

Hutto acknowledges that his view is like Goldie’s view: he calls Goldie’s *The Mess Inside* (2012) a book-length discussion of the ‘modest’ narrative proposal he himself offers (Hutto, 2016: 25). Hutto explicitly points at the explanation of reasons for action as the aim of Minimal Narrativism:

To understand which beliefs and desires were responsible for a person’s action is normally only to understand why they acted in a quite skeletal way. Maximally, to understand why someone acted requires a more or less detailed description of his or her circumstances, other propositional attitudes (hopes, fears), more basic perceptions and emotions and perhaps even his or her character, current situation and history. In short, to fully grasp why someone took action on a particular occasion requires relating that person’s ‘story’. (Hutto, 2007: 43–44)

It is less clear where Hutto stands with respect to Goldie’s view narrative sense of self. In Hutto (2016), he explicitly sets himself the aim to reject the Strong Narrativist claim that narrative is universally required for self-experience. Hutto draws from developmental psychology to support the claim that this is not the case, and that
non-pathological, non-narrative self-experience exists. Hutto points out that children usually only master narrative capacities around the age of 5 (Tomasello 2003: 276, quoted in Hutto, 2016: 38). But still, they are able to have temporally structured self-experience, since they “rely on their training and past experience to anticipate how things will generally unfold” (2016: 38). This tendency continues into adulthood:

Indeed there is every reason to suppose that the great bulk of our adult self-experience is temporally structured in innocent, child-like ways, it has its basis in embodied habits, routines and repertoires for engaging with the world formed early in our development and in later life. These ways of experiencing the world owe nothing to narrativizing and they do not cease when narrative capacities for reflecting upon and understanding the wider significance of my doings and actions are acquired. (Hutto, 2016: 38)

So, in Hutto’s view, there is no requirement to experience oneself as a being that persists through time, given that much of our self-experience is detached from the fact that we have a past, a present and a future.

That concludes the mapping of the different interpretations of the narrative claim. Next, I draw from a widely known objection to narrative theory in order to define the features of an ideal narrative account.

2.3. The anti-narrativist challenge

In this section, I answer to the main objections put forward by Galen Strawson (2004) in his landmark paper against narrative theory. By responding to Strawson from a Minimal Narrativist stance, some objections to narrativity are deflated, some are qualified and some remain a challenge.

Strawson is trying to argue that narrative theory does not accommodate the self-experience and self-understanding of people he calls ‘episodics’. According to Strawson, episodics lack all the following features:

i) Diachronic self-experience: “one naturally figures oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future” (2004: 430).
Form-finding tendency: “One must have some sort of relatively large-scale coherence-seeking, unity seeking, pattern-seeking, or most generally … form-finding tendency when it comes to one’s apprehension of one’s life, or relatively large-scale parts of one’s life” (2004: 441).

Story-telling tendency: “one must be disposed to apprehend or think of oneself and one’s life as fitting the form of some recognized narrative genre” (2004: 442).

Revision: “one will also have a tendency to engage unconsciously in invention, fiction of some sort – falsification, confabulation, revisionism—when it comes to one’s apprehension of one’s own life” (2004: 443).

Strawson’s aim is to show that any claim to the effect personhood, self-understanding or self-experience comes in the form of a narrative is wrong. He does not experience the present as embedded in the past and the future; he does not give coherence or form to the events in his life; his life does not have thematic unity; and he believes revision is damaging. Hence, insofar as narrative theory requires these four features, it is mistaken. Many narrativists focus their efforts on proving Strawson wrong. But Strawson’s criticisms are really useful to curb the excessive requirements of some narrativist views. So instead, I intend to accept his more general claim: “Some think that all normal human beings have all four of these properties. I think that some normal human beings have none of them” (Strawson, 2004: 446; emphasis added). In other words, I will accept that non-diachronic self-experience and non-narrative self-understanding may exist, and narrative theory should and can accommodate these. I go through each feature next.

2.3.1. Accepting non-diachronic self-experience

According to Strawson, narrative theory is a “descriptive thesis about the way ordinary human beings actually experience their lives” (2004: 428). According to this thesis, “human beings typically see or live or experience their lives as a narrative or story

---

15 In Against Narrativity, Strawson hints at a claim that he develops in greater detail in Strawson (2007): that narrativism is not only mistaken, but damaging. As I have said, my thesis is not concerned with moral questions, so I do not engage with that claim. Strong Narrativists would certainly have to respond to it, since they make narrativity a moral requirement—but I am rejecting Strong Narrativism in this section anyway.
of some sort, or at least a collection of stories” (2004: 428, emphases added). This corresponds with the two claims I have presented above from Schechtman’s view: that we understand and experience our lives through narrative. Strawson makes a distinction between people who are ‘episodics’, ‘diachronics’, ‘non-narratives’ and ‘narratives’ to criticise the approach.16

At the beginning of the paper, Strawson makes it very clear that ‘diachronic’ and ‘episodic’ refer to two modes of self-experience:

The first thing I want to put in place is a distinction between one’s experience of oneself when one is considering oneself principally as a human being taken as a whole, and one’s experience of oneself when one is considering oneself principally as an inner mental entity or ‘self’ of some sort. (Strawson, 2004: 429)

As seen above, according to Strawson, if one is a diachronic “one naturally figures oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future” (2004: 430). Drawing from Strong Narrativist views such as MacIntyre’s or early Schechtman’s, Strawson says that Narrativists fail to accommodate a non-diachronic type of self-experience, that he calls ‘episodic’:

“one does not figure oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future” (Strawson, 2004: 430). His main example makes clear, again, that this is a phenomenological claim. He considers Henry James, who remembered writing his early novels but said that they seemed to be written by “another person” (James, 1915: 562–3; quoted in Strawson, 2004: 430):

[James] has no doubt that he is the same human being as the author of that book, but he does not feel he is the same self or person as the author of that book. It is this phenomenon of experiencing oneself as a self that concerns me here. (Strawson, 2004: 430, emphasis added)

These examples are in striking contrast with Schechtman’s claim that ‘we experience the present in the context of a larger life–narrative’. Strawson insists that he and other episodics do not (besides James and himself, he mentions Iris Murdoch,

---

16 I have abandoned Strawson’s capitalising of these words (‘Episodes’, ‘Diachronics’, ‘Non-narratives’ and ‘Narratives’) unless I use them in a direct quote.
Virginia Woolf or Bob Dylan). Strawson’s claim is that narrativists do not make space for Episodic self-experience, where past memories have no sense of ‘mineness’ so one does not experience oneself as enduring through time.

A tempting narrativist answer is to say that Strawson just gets his own episodic self-experience wrong: that he is actually a diachronic without realising it. This is a deeply unsatisfactory answer. There are lots of people who remember their past but nonetheless feel, in the Strawsonian fashion, that those memories have no sense of ‘mineness’ (Klein and Nichols, 2012). Those people are plausibly episodics. Also, let us remember that even within narrativism there is a call for accommodating episodic self-experience: that is Hutto’s aim when he claims, as seen above, that some self-experience is “temporally structured in innocent, child-like ways” that “owe nothing to narrativizing” (2016: 38).

These examples show that rejecting episodic self-experience is not a good strategy to overcome Strawson’s challenge, if not a plainly unreasoned one. Non-diachronic self-experience, and episodics, do exist. But that does not mean that everyone is necessarily an Episodic, or that all self-experience is necessarily non-diachronic. It seems reasonable to accept, like Strawson does, that there is a distinction between two groups of people (diachronics and episodics), where that distinction is not absolute: some diachronics will be detached from some of their memories in an episodic way (2004:431), in the same way some episodics will feel a sense of mineness for some of their memories (2004:430). This idea offers grounds for rejecting both versions of Strong Narrativism, which requires persons to experience themselves as enduring beings. A plausible narrative view should then accept diachronic self-experience, but not require it universally.

2.3.2. Accepting non-narrative self-understanding

Form-finding, for Strawson, is giving coherence or unity to the events in one’s life. He explains that form-finding is a form of construal where events are not merely listed as a sequence, but given a narrative form (Strawson, 2004: 440). Matching that to the narrative claim, he means that form-finding is the task of giving meaning to events in light of each other, i.e. giving them narrative structure or the shape of a story. In distinguishing between Narratives and Non-narratives, Strawson is distinguishing between two modes of self-understanding: one based on the creation
of stories about oneself and one not based on the creation of stories about oneself. Strawson says that

self-understanding does not have to take a narrative form, even implicitly. I’m a product of my past, including my very early past, in many profoundly important respects. But it simply does not follow that self-understanding, or the best kind of self-understanding, must take a narrative form, or indeed a historical form. (Strawson, 2004: 448–449)

Diachronicity in self-experience and form-finding in self-understanding, explains Strawson, are independent of each other (Strawson, 2004: 441). This is manifest, Strawson says, by the fact that some people will try to find a narrative form in the events in their lives precisely because they are episodics (Strawson, 2004: 442). Alternatively, a diachronic may experience their life as enduring, but without formulating narrative connections among events, having then a non-narrative understanding (Strawson, 2004: 442). Form-finding “captures something that is essential to being Narrative and that goes essentially beyond being Diachronic”, and is thus both necessary and minimally sufficient for narrative self-understanding (Strawson, 2004: 443).

Strawson, then, differentiates between diachronic self-experience and narrative self-understanding, so these can be analysed separately. Strawson does not say that non-narrative understanding is the only possible mode of self-understanding. This is, again, a claim of Minimal Narrativists. A plausible narrative view should then accept that creating stories about oneself may be a route for self-understanding, but it need not be so.

2.3.3. Qualifying story-telling

Like Schechtman above (§2.2.1), I do not think much argument is needed to set aside the requirement of story-telling. Strawson discusses the authors who make the ethical requirement for thematic unity (McIntyre, Ricoeur and Taylor) with extreme suspicion, suggesting that understanding one’s life in that way is an example of Sartrean bad faith or inauthenticity (Strawson, 2004: 435). I think it suffices to say that the ethical requirement of the first version of Strong Narrativism is just too
stringent. Jeanine Weekes Schroer and Robert Schroer (who themselves defend a narrative theory of reidentification) put it well:

[A]n account of “possessing a narrative” that requires the subject to view/structure the events of her life via a single, life-long story that approximates, in quality and coherence, one written by a talented author—i.e. a life story with a gripping plot that contains no loose ends, well-positioned and developed characters, etc.—sets the bar too high. Hardly anybody meets that standard of possessing a narrative. (Schroer and Schroer, 2014: 456)

This is a good reason to reject views that require story-telling in the sense Strawson describes it. Thematic unity and adaptation of life events to a genre are only requirements for adherents to the first version of Strong Narrativism (or some of them). Hence, that cluster of views can now be definitely abandoned as a plausible framework.

However, that does not mean that story-telling should be completely discarded within a Minimal Narrativist view. All it means is that story-telling should not be a requirement. Tellingly, Strawson’s view is focused on the authors who, as seen in §2.2.1, think that a person’s narrative as a whole has the shape of a ‘quest’ for the good (the ‘recognized narrative genre’ one’s life must fit into). Firstly, since Strawson’s main complaint is that narrativists do not accommodate his own mode of self-understanding, there is no reason for not accommodating the self-understanding of those whose life does possess thematic unity —after all, Strawson accuses McIntyre and Ricoeur precisely of extracting a theory from their own experience with “fabulously misplaced confidence” (Strawson, 2004: 439). It is plausible that historical figures like Joan of Arc did see their lives as a ‘quest for the good’. Other people may have seen their lives as quests for things other than ‘the good’: Martin Luther King as a quest for freedom; Amelia Earhart as a quest for adventure; Emmeline Pankhurst as a quest for equality. Certainly, many people who see their lives as quests may be deluded, inauthentic or straightforwardly evil (insert here a historical or present tyrant of choice). But claiming that living a ‘quest for good’ is a good thing is different from claiming that such understanding of oneself

17 Hutto makes this caveat when he says that Strawson’s criticism on this issue exaggerates even the requirements of Strong Narrativism (2014: 29).
is possible; which at the same time is different from claiming that such understanding is required. Hence, that is one reason to claim that the ideal narrative view should admit that thematic unity in one’s narrative as a possible, but not required, tool for self-understanding (we cannot all be Amelia Earhart).

There is a second reason why a plausible narrative view should follow the above route. As I have mentioned, for Hutto “story-telling practices” are a generalized tool for understanding oneself and others, at least for people whose self-understanding is not pathological (Hutto, 2016: 39). More specifically:

The notion I have of myself—that is as a person exhibiting certain characteristics and fulfilling certain roles—is ... parasitic on my understanding of the selves of others ... [Our capacity for this sort of] understanding has been given to me by the stories and tales I have absorbed over the years. It is the characters I have encountered, both real and fictional, that serve as models and points of reference for my own self-understanding. I must learn the art of story-telling before I can place myself within a narrative framework, (Hutto 1997: 75, quoted in 2016: 37–38; emphases added.\(^\text{18}\)

Here, Hutto finds a place for genre within Minimal Narrativism: specific narratives which become ‘models’, which shape experience in a parallel way as genre shapes stories. Certain experiences are understood with reference to those models (think, for example, of the notion of the American dream, or the self-made businesswoman). Those stories can become models in the self-understanding of some people. And more importantly, there are stories about specific spheres of people’s lives, like motherhood, or like romantic love. Hilde Lindemann gives two examples of these models:

At some point in your life, perhaps, you might have understood your relationship with a lover by configuring it according to the stock boy–meets–girl script that structures everything from “Cinderella” to

\(^\text{18}\) The quote continues: “If we accept this then self-knowledge is not gained by looking inward (wherever this is supposed to be) but by being able to give a story or account of oneself in narrative” (Hutto, 1997: 75). This claim, which seems to require story-telling and inhabits a stance closer to Schechtman’s psychological requirement of narrativity, is in discord with Hutto’s general view. Since it is one of his earliest works on narrative and he explicitly discards the requirement of story-telling in Hutto (2016), I am going to overlook that claim.
Valentine’s Day; the story then becomes a piece of your identity—constituting autobiography. A friend betrays you, and you identify him in biblical terms: he is a Judas. (Lindemann, 2014: 4–5)¹⁹

Indeed, the landmark work on romantic love by Denis De Rougemont (1983) tracks the tropes mentioned by Lindemann, which guide many people’s notions of romantic relationships, all the way back to the story of Tristan and Isolde. More importantly, the whole feminist critique of romantic love is based on the idea that people translate models about romantic partnerships into their self-understanding and their actions. When Marilyn Friedman says that love is guided by “gender norms and ideals of romantic heterosexual love” (2003: 127), these ‘ideals’ are partly given by stories like the one described by Lindemann —she calls them “socially shared narratives” (2014:5), I come back to these in chapter 3.

With this in mind, Strawson’s objection is detrimental to Strong Narrativism but with respect to Minimal Narrativism only qualifies the boundaries within which it makes sense to think of self-understanding as having a narrative form. A view where story-telling is rejected as a tool for self-understanding would not include the self-understanding of people who live their lives with thematic unity, and would not acknowledge that stories become models in self-understanding.²⁰ So an ideal narrative theory should accept story-telling for self-understanding, but not require it.

2.3.4. Leaving revision for later

For the moment, I am going to remain silent on revision, for two reasons. First of all, despite the fact that his definition seems to carry negative normative connotations, for Strawson revision need not be wrong—it may lead, for example, to the improvement of low self-esteem (2014: 444). Therefore, a plausible narrative view could accept revision. Secondly and more importantly, narrativists do not argue for revision as a requirement, let alone in the terms put by Strawson—as ‘falsification’

¹⁹ Hilde Lindemann has published under the names of Lindemann Nelson and Lindemann, which leads to some of her work being cited under ‘Nelson’. For the sake of clarity, I use ‘Lindemann’ in all in-text references; the reference list specifies whether a work or an edition is published under ‘Lindemann Nelson’.

²⁰ For a fully-fledged theory of how not only stories understood in a generic way, but literary fiction shapes both self-understanding and moral sense, see Nussbaum (1992).
and ‘confabulation’. This is Strawson’s freestyle interpretation of the consequences of the claim that we are the creators of self-narratives. Just to be clear: no narrativist says that we should confabulate accounts about our lives. So considering revision as a feature of narrative theory instead of a consequence of the claims in narrative theory is just wrong.

Admittedly, revision is a potentially problematic consequence of narrative theory. If it is true that we create narratives about ourselves, the risk for confabulation is there. Imagine, for example, a doctor who administers involuntary euthanasia to her dying patients for years. Her story, to herself, is that of a compassionate saviour of the sufferers. We would morally, and certainly legally, object to such an interpretation: her story is that of a murderer. I will deal with this problematic consequence of narrativity in due course (chapter 3). But before dealing with the consequences of a narrative view, it is necessary to establish the features of such a narrative view. I then ignore, for the time being, this potential consequence of narrativity, which is miscategorized by Strawson as a feature of narrative theory.

2.4. The challenge to Minimal Narrativism

Let us sum up the responses to Strawson’s challenges so far. Through the evaluation of Strawson’s anti-narrativist approach, I have shown that a plausible narrative theory should follow the following route: in self-experience, diachronicity (experiencing one’s present as a being embedded in one’s past and future) is possible, but not necessary. In self-understanding, form-finding (creating stories about oneself) and story-telling (understanding one’s life or parts of one life in terms of genre) are also possible, but not necessary. How do Hutto’s and Goldie’s views differ from this plausible narrative theory?

According to the authors themselves, their views do not differ from this narrative theory at all. Hutto’s position on the non-necessary nature of diachronicity, form-finding and story-telling has been repeatedly shown throughout the chapter. Yes, appealing to people’s stories is the route to discover their reasons for action, but that does not mean that the creation of stories is a requirement for self-understanding. First of all, Hutto’s notion of story-telling merely requires engaging in narrative practices (i.e. telling and listening to stories), whereas Strawson understands story-telling as the encircling of events within a literary genre. Hutto’s lengthy rejection of
Strong Narrativism in his (Hutto, 20016) suffices to show that he does not place such a requirement on story-telling. Secondly, he explicitly says that episodic self-experience is non-pathological, and non-narrative self-understanding is non-pathological too. As seen above, Hutto accepts that there are everyday experiences which do not take narrative form. We have also seen that Hutto states that listening to stories and applying them as models is the “normal route for acquiring the competence needed for making sense of ourselves and others in terms of reasons” (Hutto 2016: 39). Does this mean that ‘normal’ self-understanding is narrative? Not necessarily: it is relevant to note here that Hutto’s view is framed within developmental psychology and, when he states that it is the normal route of acquisition of competence for understanding reasons, it can be interpreted as the ‘normal’ route followed during childhood, not the ‘normal’ way of thinking about oneself at all times. “I must learn the art of story-telling before I can place myself within a narrative framework”, he says (Hutto 1997: 75). But he does not require anywhere that I must place myself within such a network. Hutto seems to pass the requirements for a plausible narrative view.

Let us turn to Goldie. Again, he explicitly and vehemently rejects the Strong Narrativist demands for living one’s life as a story. He rejects the requirement of diachronic self-experience and of story-telling. Regarding form-finding, we have seen that Goldie not only does not consider it necessary but actually warns against the tendency towards form-finding at all times (2012: 161–171). He does not even believe that narrative thinking always consists in creating stories, since as we have seen ‘narrative thinking through’ can take other forms like experiential imagining. So, again, Goldie’s theory seems to match the conditions for a plausible narrative view.

This is not all there is to it, however, because Minimal Narrativism as posited by Hutto and Goldie still faces an objection that I have not yet mentioned: the problem of triviality. Strawson formulates the objection in the following way:

If someone says, as some do, that making coffee is a narrative that involves Narrativity, because you have to think ahead, do things in the right order, and so on, and that everyday life involves many such narratives, then I take it the claim is trivial. (Strawson, 2004: 437–438).
It is not clear to whom this criticism is directed (plausibly against David Velleman, whose narrative account of agency in his (Velleman, 2006) makes some remarks on coffee-making). Some narrativists have tried to save narrative theory from this objection by pointing out that with this example Strawson is admitting that coffee-making is a narrative which is irremediably embedded in wider narratives (Rudd (2012: 184–184); Schroer and Schoer (2014: 455)).

I think it is more illuminating to take the objection in its strongest form: if a narrative view does not require form-finding, i.e. it neither requires relating a series of events, nor putting these events in the shape of a story, then narrative theory is trivial. Lamarque, who is another prominent anti-narrativist, has more effectively directed the triviality objection towards Minimal Narrativism, specifically Goldie’s view:

Perhaps [Goldie’s] modest view can accommodate the dispositions of the likes of Strawson and me. After all, narrative thinking does not imply that full-blown, articulate narratives govern our every act or thought, and certainly not the need for any overarching life–narrative; narrative thinking can be indeterminate, lacking beginnings and ends, vague through and through. And a narrative sense of self really involves just an awareness of past, present, and future. But along this path it might seem that the appeal to narrative is trivial and insignificant. Maybe a narrative is just a thought that conjoins two descriptions or sentences. Isn’t ‘I missed the bus so I was late’ a mini-narrative? And who could deny that thoughts of that kind are commonplace? The dilemma is that if Goldie emphasizes the trivial end of narrative then he might see off the worries from the sceptics but his claims about narrative seem weak. But if he insists on something more substantial then he fails to confront the fact that there are those of us who don’t have strong narrative dispositions. (Lamarque, 2016: 58–59)

Lamarque believes that the term ‘narrative’ refers to “a complete, rounded story with a beginning, middle and end that helps make sense of complex events” (2004: 405). So certainly a ‘conjunction of two sentences’ does not qualify as a narrative. In the quote above, Lamarque is endorsing the idea that I have put forward in this section: that Minimal Narrativism accommodates episodic self-experience and non-narrative self-understanding (entailed by his mention of ‘the likes of Strawson and me’). But
he is implying is that a narrative view that does not require form-finding does not say anything interesting about how people understand themselves.

There are two ways to avoid the challenge of triviality. The first one is to require form-finding for self-understanding and reject non-narrative self-understanding. As I have shown, that would be an implausible narrative theory in light of Strawson’s objections. The second one is to show that a narrative view which does not require form-finding for self-understanding is not trivial. That is the route I will follow. I believe that there is a piece missing in the discussion for that strategy to succeed. In the next chapter, I develop a non-trivial Minimal Narrativist view on the self-concept.

**Conclusion**

Throughout this chapter, I have put aside several approaches in order to lay out the foundations for my theory on what being in love is. I have departed from the philosophy of emotion given its focus on justification, and from personal identity given its focus on autonomy. Instead, I have placed my theory at the level of the self-concept: the beliefs that a person has about oneself. Some of these form a person’s characteristics, which are like Helm’s conception of values: descriptions of one’s life according to which one’s life is worth living.

I have shown that in order to have a plausible account of the narrative self-concept, Strong Narrativism should be abandoned. In light of Strawson’s criticisms, a plausible narrative theory should not require diachronic self-experience, form-finding in self-understanding or story-telling, understood as one’s self-understanding adapting to the form of a genre. Minimal Narrativism fulfils those requirements, but is in turn open to Strawson’s objection of triviality: if Minimal Narrativists do not require form-finding, why would narrative be relevant at all? In the next chapter, I offer an account which avoids that objection.
Chapter 3

The narrative self-concept

In this chapter, I lay out a view on the narrative self-concept which will be the foundation for my account of being in love. Due to the anti-narrativist challenge, a non-trivial account of the narrative self-concept is needed. Minimal Narrativist views have several gaps that need to be filled in order to make such an account available for defining what being in love is: that is the aim of this chapter.

I start in §3.1 by developing the role of narrative in Minimal Narrativism in more detail. I explain that narratives are a route to maximal understanding, which is the understanding of a person’s reasons beyond mere desires and beliefs. I explore Hutto’s and Goldie’s distinct but complementary explanations of why narratives provide with maximal understanding: the fact that reasons for action are partly acquired through narrative models (Hutto) and the fact that narratives have higher explanatory power than merely causal explanations due to their meaningfulness and import (Goldie).

In §3.2 I show that there is a missing piece in Minimal Narrativism. Namely, as currently formulated Minimal Narrativism does not fully account for the fact that narratives do not only contribute to self-understanding, but also to the understanding of the world. I argue that Minimal Narrativism underdevelops the role of salience, which determines how events are selected and interpreted.

In §3.3 I argue that salience has a crucial role in a person’s self-concept. I explain that there is a feedback loop between self-concept and action. This feedback loop is grounded in meaningfulness and salience: only by acknowledging that it can be claimed that narratives are a route for maximal understanding. On this basis, the role of narrative in the self-concept is non-trivial.

In §3.4 I acknowledge that an anti-narrativist can still argue that the feedback loop between self-concept and action need not be described through form-finding. Following Mary Jean Walker (2018), I explain that the self-concept is necessarily diachronic, given that present traits of the self-concept are necessarily defined in relation with past self-concepts. On this basis, the role of narrative is non-trivial for the self-concept, even if reasons for action are understood in a non-narrative way.
Finally, in §3.5 I briefly deal with two objections coming from truth and revision. Firstly, Strawson and Lamarque have expressed that understanding persons through narratives is dangerous, given the damage that false narratives can make. I explain that this is an objection to a narrative theory that aims at true self-knowledge, not understanding in general. Secondly, I return to Strawson’s objection of revision, which considers it dangerous that people can re-formulate narratives about themselves. I argue that such an approach ignores the fact that self-concepts are formed relationally. For that reason, I argue that revision can be as detrimental as it can be beneficial – as well as being one of the key features of my view on being in love, which I will develop further in chapters 4 and 5.

3.1. The role of narrative in Minimal Narrativism

The anti-narrativist objection of triviality assumes that a view on the narrative self-concept needs to prescribe form-finding (i.e. creating stories about oneself) in order to not be trivial. That assumption, however, fails to fully account for the role that narrative has in Minimal Narrativism, which I define in this section.

3.1.1. Maximal understanding

Given that it does not require diachronic self-experience, Minimal Narrativism is mainly concerned with narrative self-understanding. Hutto and Goldie coincide in pointing at the explanatory power of narratives:

Maximally, to understand why someone acted requires a more or less detailed description of his or her circumstances, other propositional attitudes (hopes, fears), more basic perceptions and emotions and perhaps even his or her character, current situation and history. In short, to fully grasp why someone took action on a particular occasion requires relating that person’s ‘story’. (Hutto, 2007: 43–44)

Narratives often provide explanations of why someone had a particular motive, or why someone has a particular character or personality trait, or why someone was drunk, depressed, or angry. And the explanations that we get are narrative–historical explanations: they locate the motive, the trait, the undue influence on thinking, within a wider nexus, in a way that enables us to understand more deeply why someone did the thing.
that they did through appeal to aspects of their personal history or circumstances. (Goldie, 2012: 20)

Hutto and Goldie coincide: narratives can potentially explain action (which is not to say that all narratives explain; see §2.2.3). And they can do it better than mere beliefs and desires, given that “[t]o understand which beliefs and desires were responsible for a person’s action is normally only to understand why they acted in a quite skeletal way” (Hutto, 2008: 43). Narratives explain ‘maximally’, to use Hutto’s term. Hence, I take that in Minimal Narrativism the role of narrative is maximal understanding of people, specifically of reasons for action. Let us remember that, in light of the anti-narrative challenge, narrative has to be necessary for self-understanding without the requirement of form-finding. My strategy is to show that narrative is necessary for a specific kind of understanding: maximal understanding. If it can be proven that narrative is necessary for maximal understanding, but not for self-understanding in general, then narrative theory is not trivial. Before I offer the Minimal Narrativist argument for that conclusion, I make two brief observations.

Firstly, it may be noted that my use of the term ‘maximal understanding’ seems to be wrong, and I should instead use ‘maximal self-understanding’. This is not an oversight. Minimal Narrativists, and very explicitly Goldie, claim that narratives are not only a route to understanding oneself but the world in general:

[M]emories of our past, our deliberation and plan-making about our future, infuse our understanding of ourselves, and of others, in the present (Goldie, 2012: 116; emphasis added)

Our narrative sense of self is present to us not only when explicitly thinking of our past and future or when explicitly engaged in narrative thinking. It is also intricately involved in the way we engage with and think of our present environment and of ourselves and other people. This kind of awareness, more or less conscious, more or less something in which one is actively engaged, shapes the content of our thoughts and feelings about ourselves, others, and the world. (Goldie, 2012: 119; emphasis added).
As I will eventually show, while Minimal Narrativists offer a convincing case for the maximal explanatory power of narratives in self-understanding, they fall short of backing the role of narrative in the understanding of other people and the world. This is crucial for my thesis, given that my argument on being in love as a trait of the self-concept rests on this role. For the moment, I am going to focus on the Minimal Narrativists’ claims about self-understanding and then move to my critique in §3.2.

Secondly, Hutto has his own answer to the triviality challenge. Hutto explicitly says that he is interested not in what are reasons for action, but in how we acquire the ability to find reasons for action: “how we become skilled at the practice of predicting, explaining and explicating actions by appeal to reasons” (Hutto, 2007: 43). Drawing from empirical studies, he endorses the hypothesis that children learn to understand reasons for action through engagement with stories. That is the basic claim of his theory, which he calls the Narrative Practice Hypothesis (Hutto, 2007: 53). The role of narrative, for Hutto, is not only to provide with reasons for specific actions. It is also the route to acquire the ability to understand people’s actions, in general, in terms of reasons. On that basis, Hutto says, narrative is not trivial. Hutto is right: if people learn to understand actions in terms of reasons through narrative during childhood, then that is an argument against the triviality objection. But it is not the argument I am concerned with here. I am trying to arrive at a view on the narrative self-concept where form-finding is not required, but narrative is nonetheless non-trivial. Hutto’s conclusion on the triviality challenge is interesting, but is tangential to my aim – I will then not engage with Hutto’s response to Strawson.

3.1.2. Narrative explanation vs causal explanation

So how do narratives contribute to maximal understanding? I start by giving a brief general response. A narrative product is the relation of a series of events with a specific form, which differentiates them from purely causal explanations. The type of belief/desire-based understanding of reasons for action that Hutto calls ‘skeletal’ is formulated in causal form: ‘I left the party because I wanted to sleep’; ‘I didn’t believe you wanted cake so I didn’t offer you’. In narrative explanation, a causally organised chain of events is not enough to convey the kind of intelligibility that is required for self-understanding – to “making sense of ourselves and others” as Hutto puts it (2016: 39). Schroer and Schroer have a good example of how a narrative
relation of events differs from a purely causal relation of those events. They bring in two individuals, Perry and Larry, who both have the desire of making a cup of coffee. When prompted for an explanation of why he is making a cup of coffee, Perry gives several replies:

[He] will say things like “I make coffee because I want some” or “First, I grind the beans. Second, I boil the water…” or “I hear the water boiling, I smell the coffee, I feel the warm cup in my hand, etc.” (Schroer and Schroer, 2014: 456).

Larry, on the other hand, has something more to say about the process of making coffee than just going through the motions:

My desire for coffee is part of my overall addictive personality… I also really desire cigarettes and cheap booze. This makes the experience of making coffee somewhat unpleasant for me, for it reminds me of this overall flaw in my character. Making coffee is a daily reminder, in a small way, of both my past struggles with addiction and of the kind of person I am striving to be—the kind of person who doesn’t need chemicals to feel good about himself. (Schroer and Schroer, 2014: 456)

That is a narrative explanation of Larry’s coffee-making experience. Larry’s explanation is narrative, according to Schroer and Schroer, because it approximates the logic of a story and involves self-reflection on how past events have shaped Larry’s life (and probably will continue doing so). It does not need to have the quality of a story written by a professional, and it does not need to be made explicit—-but it would be accessible if Larry was prompted to speak or think about the experience of making coffee. The difference between Perry and Larry is that for the latter, the coffee-making is embedded in the events that have happened in his past and a projection of what may happen in his future. Re-writing Schroer and Schroer’s example in a purely causal way may shed some more light on the explanatory capacity of narratives. Imagine this is what Harry says about his coffee-making experience:

In the past, I was addicted to certain substances. That past addiction has caused my present addiction to certain substances. I do not like my present addiction to certain substances. Coffee is one of the substances
I am addicted to. This fact, together with my past addiction which continues into the present, causes an unpleasant feeling. Making coffee elicits past memories of consumption of certain substances I am addicted to. The unpleasant feeling causes a desire to avoid future episodes of consumption of certain substances.

There seems to be a difference in explanatory power between Perry’s belief/desire-based, Larry’s narrative-based and Harry’s causal-based accounts. This is because of the nature of narrative products as a result of a narrative process. The process of creating a narrative is, at the same time, a process of selection and a process of interpretation. That is, narrating involves the choice of certain events which will have a specific place in the narrative (selection). This is also done in causal accounts. In narrative explanations, however, the events related acquire certain meanings in according to how they are presented: what place they occupy in the narrative, with which other events they are connected and in which way (interpretation). This is what Goldie calls emplotment (a term he borrows from Paul Ricoeur):

The process of emplotment involves shaping, organizing, and colouring the raw material into a narrative structure. Shaping involves selecting raw material with the appropriate degree of richness, and shaping it in a way that is appropriate to the narrative. Organizing involves configuring the raw material into a narrative, with a beginning, a middle, and an end. And colouring (not a necessary ingredient of emplotment, perhaps, but a typical one) involves bestowing evaluative and emotional import to what might otherwise be a bare description of what happened. (Goldie, 2012:10–11, emphases in original)

The relation of Perry’s beliefs and desires does convey intelligibility about Perry’s actions; so does the relation of Harry’s causal account. However, with Larry’s interpretation and shaping of the events, we get access to the meaning that these events have for Larry. That gives narrative-based explanations a higher explanatory power and ultimately places narrative as a route for maximal understanding. I move on next to describe how Hutto and Goldie justify the higher explanatory power of narratives.
3.1.3. Hutto’s narrative models

Hutto claims that in order to understand reasons for action—our own and others’—we need to look at people’s stories. Narratives are then a route to understand oneself and others, since it is through narratives that we understand why we do what we do. But, Hutto warns, narratives are not necessary for understanding: some actions, like those rooted in habit, need not be explained through narrative. Also, although narrative understanding is prevalent among non-pathological population, that does not mean that there cannot be non-pathological, non-narrative understanding—like Strawson’s; Hutto puts the onus on him to show that non-narrative understanding is not anecdotal. These claims summarise the description of Hutto’s view I have offered in the previous chapter.

Because Hutto’s is mostly concerned with the role of narrative in our ability to understand actions in terms of reasons (see §3.1.2), his argument is focused on defending that claim, and not in developing in detail why narratives have a higher explanatory power to understand reasons for action. Hutto calls for looking at a person’s story to understand these reasons (Hutto, 2008: 43–44). For Hutto, narratives provide with ways to think “about the particularities of our characters and our situations” (2016: 25). Narratives are also a way of reflecting about and adjusting “personal habits, dispositions, styles of engagement and attitudes” (2016: 25). Ultimately, Hutto does not offer full support for these claims, which he believes are justified by his Narrative Practice Hypothesis—the hypothesis that we acquire the ability to understand actions in terms of reasons through narratives. However, we can still extract one specific claim on self-understanding from Hutto’s account: the claim on narrative models.

Hutto explains narratives models, i.e. the stories we access about how certain events in life are interpreted, become points of reference in self-understanding. Hutto elaborates further on the implications of narrative becoming models:

By providing a window on our actual and possible doings the narratives we fashion about our lives can help us to decide, for example, if our taking this or that action is something we want to figure in the story or stories that we would want told about who we are. It would be hard to deny that, personal narratives are at least one natural means—and
possibly, even the primary or ideal means—for understanding and shaping ourselves (Hutto, 2016: 26).

This is basically all Hutto says about narrative models, so I have to do some interpretive work. Here Hutto seems to say that narratives not only explain with reasons for action, but they actually provide with reasons for action as well. For example, having grown up listening to the stories of the early 20th century adventurers, a country girl who is deciding which degree to choose faces a dilemma: does she continue the family tradition and become a lawyer focused on agricultural matters, or does she go abroad and study a degree that will allow her to travel the world? When making a choice, the country girl has more than a series of isolated actions in front of her: she is thinking about ‘the story she would want to be told about her’. That is how I interpret Hutto’s narrative models, so the story, for example, of Amelia Earhart, becomes a source of reasons for action for the country girl who read about Earhart when she was little. I think this is what Hutto means, with less grandiose narratives models also shaping everyday life. ‘Do I want this to be the day where I start living a healthier lifestyle?’, one may think as the future narrative of one given date. Just to insist, Hutto does not believe all actions are explained through narrative, so not all actions need to be sourced from narrative in this way either.

One need not be conscious of the narrative models which are points of reference for one’s actions. That makes sense in the case of the country girl, who probably will not be explicitly applying Amelia Earhart’s story to her future decision. However, this story that frames her thoughts about her current dilemma, and subsequent actions. Maximal understanding of the country girl would require not only looking at her current beliefs and desires, but also about how these beliefs and desires were formed—in Hutto’s account, at least partly through narrative models. Bringing this to the examples above, we can see that Larry is reproducing the narrative model of addiction, both in his actions and in his account of those actions. This is something which we cannot see in Perry’s belief/desire–based account, and which is obscured in Harry’s causal account. This is one sense in which narratives can be a route for maximal understanding.

Nevertheless, narrative models are admittedly in the furthest end of maximal understanding. Explaining someone’s actions by directly appealing to narrative
models would plausibly be an over-psychoanalisation of why a person did what they did at a given moment. Although narrative models will ultimately be crucial to my account, they do not by themselves provide a convincing argument for the role of narrative in maximal understanding. Hutto may not need that argument given his aim, but I do. Luckily, Goldie’s account fills in the gaps and offers a more plausible account on how narratives contribute to maximal understanding, without the need to go as far from a person’s actions as the narrative models which are a reference for such actions.

3.1.4. Goldie’s meaningfulness and import

Before embarking on a detailed description of Goldie’s account, it is worth noting that, like Hutto, Goldie has a specific aim in his discussion of narrative thinking. Goldie’s main aim is to explain the mechanisms underlying the process of narrative thinking.

Largely, Goldie’s discussion on narrative thinking responds to what he calls ‘the problem of the audience’: since narratives require a narrator and an audience, how can that work when one thinks about oneself? (2012: 40). Goldie’s solution is that a person’s current perspective is both internal (as belonging to the subject doing the thinking and formulating the thoughts) and external (as belonging to the narrator thinking about herself as a character in those thoughts) (2012: 26). A person can inhabit both perspectives at the same time when thinking about herself in the past and/or the future. This idea is best understood with an example. Goldie invites us to consider the following narrative of a tennis doubles match: “My partner and I were both chasing the same lob, and then we hilariously fell over each other in a tangle” (2012: 38). This is how he explains internal and external perspective:

In relating this narrative, I tell how at the time I was highly amused. This emotional response is internal to the narrative in just the sense that it is I in the narrative who felt amused— I as the internal character at the time found the incident to have this particular internal emotional import. Now, as external narrator I still consider what happened to be amusing, and this is expressed in my narrative (Goldie, 2012: 38)
That is, I am in one sense external to a narrative when I am telling or thinking a narrative through. A person can produce a narrative about a series of events and describe them by pointing at that person’s own evaluation of these events—in the example, as amusing. Sometimes, the narrative I am telling or thinking through is about myself. In that case, I am internal to the narrative, and thus inhabit an internal perspective, when I am a protagonist of the narrative. To further help the grasping of the two perspectives as inhabited by the same person, think about a conference dinner where you feel elated, like you are the soul of the party. The following day, however, you tell your friend: “Yesterday I was elated and feeling like I was the soul of the party, but now I realise I was merely drunk. I am mortified”. You, as a narrator of the dinner (external perspective), are mortified now; but you, as the character in the dinner (internal perspective), are elated in the narrative (which happened in the past).

This is a very interesting contribution on how narrative thinking works on a technical level, which Goldie also backs with evidence from developmental psychology (2012:29–30). However, and although acknowledging the distinction between internal and external perspective is necessary to fully understand my discussions of Goldie’s account in this chapter, it is not directly relevant for my purposes. What I am trying to locate is the role of narrative in understanding, specifically the role of narrative products (i.e. stories about oneself and others). Goldie’s argument in that sense is located in his account of meaningfulness, i.e. the feature of narrative that gives them the status of explanations.

3.1.4.1. Internal meaningfulness

According to Goldie’s definition of internal meaningfulness,

\[ \text{a narrative can be meaningful by revealing how the thoughts, feelings, and actions of those people who are internal to the narrative could have} \]

\[ \text{21 For that reason, I omit altogether the notions of dramatic irony and free indirect speech, which according to Goldie enable the inhabiting of an internal and an external perspective at the same time in narrative thinking about oneself. He develops the concepts throughout several chapters in his book, mainly in Goldie (2012: 27-56). See also Goldie (2007) for a paper-length discussion on dramatic irony and mind-reading. For an objection to the role Goldie gives to dramatic irony, free indirect speech and internal and external perspective in autobiographical thinking, see Lamarque (2016: 50-56).} \]
made sense of them from their perspective at the time—that is, from their internal perspective. (Goldie, 2012: 18)

Goldie provides the following example:

Mrs Jones went to her first meeting with her Harley Street consultant. She felt overawed because of his domineering manner. She decided that it was best to say nothing for fear of saying something stupid and ignorant (Goldie, 2012:18).

Goldie explains that from this narrative, we—the audience—can appreciate which psychological states Mrs Jones is in: “we might simply understand that Mrs Jones felt overawed and so on” (2012: 18, emphasis added). Sometimes, the narrative may appeal to a person’s motive: by stating the reasons someone had to do what they did, we can understand what they did and what was in their minds by filling in the gaps (Goldie, 2012: 19). For example, narratives can appeal to “a character trait of which that action is an expression” (2012: 20). In the narrative above, we could infer that Mrs Jones is not very assertive. In other cases, narratives can also appeal to what Goldie calls “influencing factors” which are not directly psychological states but influence these, like being drunk or having a cold (2012: 20). These three can come together, as seen in §3.1.1 (“Narratives often provide explanations of why someone had a particular motive, or why someone has a particular character or personality trait, or why someone was drunk, depressed, or angry” (Goldie, 2012: 20)).

Meaningfulness, then, is the ability to explain; specifically, the ability to provide with maximal understanding. Given that narratives can reveal actions, thoughts and feelings, narratives are a route to maximal understanding: understanding beyond beliefs and desires. However, as I showed with the examples of Perry, Larry, and Harry, explanations of actions, thoughts and feelings need not take a narrative form. However, without narrative form they do not reveal the evaluative and emotional import that actions, thoughts and feelings have for the people involved in them.

3.1.4.2. Evaluative and emotional import

We saw that Hutto limits himself to say that narratives are a route to explain people’s reasons for action Goldie digs deeper into the psychological grounds for such reasons:
Things matter to people, and a narrative involving people can capture the way things matter to them. We already have this in the story of Mrs Jones: she felt overawed because of her consultant’s domineering manner, and she thought it best to say nothing for fear of saying something stupid and ignorant. The internal meaningfulness of Mrs Jones’s behaviour is thus bound up with the evaluative and emotional import that the situation had for her: her evaluation of her consultant as being domineering, and her fearful response (Goldie, 2012: 23).

Since internal meaningfulness is an explanation of action, what Goldie is saying here is that in order to understand Mrs Jones’s actions in the situation we need to look at how she judged such situation and how she felt about it. And narratives, according to Goldie, show precisely that. Mrs Jones “feel overawed” (that is how she felt); “because of her consultant’s domineering manner” (that is the motivation for the feeling); “she thought it best to say nothing” (that is her choice of action); “for fear” (that is her emotionally motivated reason for choosing her action, i.e. saying nothing); “of saying something stupid and ignorant” (that is her evaluation of how any contribution from her would be). Here is where we can see clearly that an explanation with emotional import does have more explanatory power that one without it. Imagine that instead of Goldie’s narrative, we had the following explanation of how the meeting went:

Mrs Jones went to see her Harley Street consultant and did not say a word during the meeting because of the consultant’s attitude.

Surely, we have some understanding of what happened. But we have much a better understanding when we see the evaluation that Mrs Jones, as a protagonist of the narrative, made of the consultant’s attitude, how it made her feel, and how she felt about herself. The initial narrative shows how evaluative and emotional import is connected to internal meaningfulness: judgements and emotions related to an event are factors that contribute to explain how people act in that event. Since narratives reveal these judgements and emotions, they have a fundamental role in understanding person insofar they have greater explanatory power than causal explanations. Bringing back the earlier examples, now we can see the difference between Larry’s and Harry’s explanations: Larry’s has emotional and evaluative import. Like in the case of Mrs Jones’s evaluation of the consultant, in Larry’s
account we have access to the emotional and evaluative import the situation has for him: “This makes the experience of making coffee somewhat unpleasant for me, for it reminds me of this overall flaw in my character”. However, this is not the only way narratives have higher explanatory power than causal accounts. Given that, in narrative thinking (i.e. self-understanding), the person doing the thinking is both character and narrator, narratives also explain why we understand ourselves in certain terms. This is explained by Goldie through external meaningfulness.

3.1.4.3. External meaningfulness

Goldie adds a second way in which narratives can explain. He calls it ‘external meaningfulness’:

a narrative can be meaningful by revealing the narrator’s external perspective: his or her thoughts and feelings that throw light on why the narrative was related (or just thought through) in that particular way (Goldie, 2012: 17).

We saw in §3.1.2 that narratives are both selective and interpretative. External meaningfulness is about the form the narrative takes: it is an explanation of why certain events were selected, and why they were interpreted in a certain way. Then, external meaningfulness is also the potential to offer an explanation: not about actions, but about how that action is understood for the person describing it (i.e. the narrator). I show now how external meaningfulness is related to understanding oneself and others.

In narrative thinking about oneself, the narrator and the audience are the same person: the person who is doing the thinking. If through internal meaningfulness narratives reveal a person’s reasons for action in an event, through external meaningfulness narratives explain why a person tells herself how that event went (including her reasons for action) in a specific way. Mrs Jones may arrive home, remember the meeting and think about what has happened in the following way:

I went to the first meeting with my Harley Street consultant. I felt overawed because of his domineering manner. I decided that it was best to say nothing for fear of saying something stupid and ignorant.
The reason the shape of the narrative matters becomes apparent when we see that Mrs Jones could have recounted the event in a very different way. Mrs Jones could have gone to the meeting, have made the same judgements, have had the same emotions and have acted in the same way and, nevertheless, formulate a different narrative about the event. For example:

I went to the first meeting with my Harley Street consultant. He was one of those domineering people that I have repeatedly promised myself not to allow to scare me. But I did it all over again, like I always do, and said nothing out of fear.

Both of these narratives may be thought through to oneself, but they could also be told to another person in the same shape. The interesting question here is how the shape of the narrative reveals something about Mrs Jones (i.e. is a route to a maximal understanding of Mrs Jones). Mrs Jones focuses on her decision to say nothing in the first narrative, but on her (broken) promise to herself in the second. Likewise, in the previous example, Larry embeds the episode of making coffee in a background of addiction, and not, say, as the start of a day full of energy; he also tells it with a pessimistic tone.

The narratives we formulate about the events in our lives can explain why we do certain things, as well as what we are thinking and feeling while we do them. But also, they can offer understanding about the way we think about these actions and mental states. In turn, the way we think about our situation will influence our deliberation and choices about what to do. With this sort of meta–self-understanding Goldie goes a step further than Hutto in showing how narratives can offer maximal understanding. We can understand not only what we did, but also how we conceive what we did, which has a crucial role in our actions. It now becomes clear why we have a maximal understanding of Larry, but not of Perry.

### 3.1.5. Summing up and setting an objection

In light of the above, Minimal Narrativism opens a route to avoid the objection of triviality. Namely, narratives about ourselves are necessary for maximal understanding. With a mere causal account, we cannot access the explanations narrative makes available: why a person acts and why a person construes their
actions in a certain way (which is grounded both in Goldie’s external meaningfulness and Hutto’s narrative models).

Recall that Minimal Narrativism does not require the creation of narratives for understanding of oneself or others. Minimal Narrativists do not claim that we cannot access any understanding whatsoever of people’s actions if we do not look at their narratives. If Larry is he is making coffee because he wants to drink something, I will still have understanding of his reasons in terms of desires. This would not be maximal understanding, but would nevertheless be understanding. Equally, we do not always need maximal understanding. Larry’s narrative was illustrative of how narratives can maximally explain a coffee-making episode, but it would be bizarre if Larry offered that narrative every time he got asked why is he making coffee.

However, in some cases, we do need to understand ourselves and others maximally. We aim for maximal understanding when we go to therapy, when we try to find the root for our desires and beliefs. Mrs Jones may be oblivious to the fact that her leaving the consultant meeting without saying a word is rooted in her inferiority complex. But Larry, for example, is aware that the stress that making coffee provokes is embedded in his history of addiction. Mrs Jones has a minimal understanding of his own actions, while Larry has a maximal understanding of his own actions. And if Mrs Jones wants to stop being silent in the face of domineering characters, maybe she will need to understand why she stays silent in the first place: in that case, she will need maximal understanding. What Minimal Narrativists show is not that belief/desire–based explanations of actions are wrong, but that they are not the only route for explanations, and that sometimes they are not sufficiently explanatory.

I now note a flaw in Minimal Narrativism. I have accepted Strawson’s claim that some people may not engage in form-finding in self-understanding (i.e. may not formulate narratives about themselves). For these non-narrative people, if they exist, the above account of maximal understanding is still trivial, given that if Strawson is right they would be able to understand themselves maximally through causal accounts similar to Harry’s. Is narrative still non-trivial for the self-concept in these cases?
Strawson or Lamarque could still argue that Larry’s and Mrs Jones’s accounts are merely causal accounts, and the notion of narrative is not doing any work. Such a response would overlook the well-argued claim that narratives are a route to maximal understanding. However, maximal understanding greatly rests on the evaluative and emotional import of narratives. The anti-narrativist could say that evaluative and emotional import are just extra causes in a causal account, and that you do not need a narrative to incorporate these. I will accept the strongest formulation of the anti-narrativist challenge and find a way to show that narrative is necessary for the self-concept, even for people who do not engage in form-finding when understanding themselves.

3.2. The missing piece in Minimal Narrativism

3.2.1. Understanding the world

Minimal Narrativists explain how narratives contribute to maximal understanding of persons, but fall short of their wider aim. Larry and Mrs Jones can have a better understanding of themselves through their narratives than through causal accounts, and we can have a better understanding of others (Larry and Mrs Jones) by looking at their narratives. However, let us remember what Goldie believes narratives can explain: the narrative sense of self “shapes the content of our thoughts and feelings about ourselves, others, and the world” (Goldie, 2012: 119; emphasis added). Nothing in the above account really justifies the claim that narratives shape our understanding of the world. I believe, however, that the claim is correct, and give my argument for it next—which eventually will result in a non-trivial narrative account of the self-concept.

According to Minimal Narrativism, narratives contribute to the maximal understanding persons by their ability to reveal different features of a situation (people’s actions, people’s judgements and emotions related to those actions, people’s understanding of how the action unfolded). Another way of saying this is that Minimal Narrativism, as currently formulated, is grounded in meaningfulness: narrative contributes to understanding in virtue of its ability to provide with an explanation. In order to give an explanation, they need to be narrative products, which as seen above can be rejected by anti-narrativists as mere causal accounts.
However, understanding should not be grounded in meaningfulness only—it is not only explanations that matter.

Goldie implicitly discusses a person’s understanding of the world as follows. Goldie rightly acknowledges that when we try to explain a given event, there are many multiple causes that can be cited. The causes that one cites will depend on one’s interests (Goldie, 2012: 15). Goldie brings in David Lewis’s (1986) discussion on multiplicity of causes to illustrate why different narrators establish different causal connections in the face of a chain of events. Here is what Lewis says about the possible causes people will cite for a car crash involving a drunk driver and a bald tyre:

If someone says that the bald tyre was the cause of the crash, another says that the driver’s drunkenness was the cause, and still another says that the cause was the bad upbringing which made him so reckless, I do not think any of them disagree with me when I say that the causal history includes all three. They disagree only about which part of the causal history is most salient for the purpose of some particular inquiry. They may be looking for the most remarkable part, the most remediable or blameworthy part, the least obvious of the discoverable parts. (Lewis 1986: 215 in Goldie 2012: 16)

In this respect, Goldie says that narrative accounts are similar to causal accounts. That is because the elements of the narrative depend on the narrator’s interests:

Similarly, in narrative explanations, the elements that the narrative mentions will depend on our interests. I might say that I missed the train because my partner refused to give me a lift in the car (she was too busy, she said), whereas she might say that I missed the train because I dawdled so much that walking to the station would not get me there on time (and she really was too busy to give me a lift). (Goldie, 2012: 16)

Goldie does not elaborate further on what he means by ‘interests’. The notion of interests is a very open-ended one, which can be understood as aims or desires. It can also be understood in terms of flourishing: an agent’s interests are the things that will promote her well-being. In the quotes and examples he provides with, Goldie
gives us two clues. Firstly, Lewis understands interests as the narrator’s communicative goals. That is what he means with salience being directed to “the purpose of some particular inquiry”. For example, if the narrator is trying to communicate about responsibility in the case of the car crash, they will try to find a causal link which can be subject to blame. An inanimate object – the tyre – cannot, while a person – the driver – can. On the other hand, if they are trying to emphasise the randomness of life and death, they may point to some non-agential force – the tyre. Secondly, Goldie may seem to understand interests in terms of desires. In his example of missing his train, his main interest is a desire to arrive on time, while his wife’s main interest is the desire to finish the tasks she was busy fulfiling. So, given that communicative goals could also be broadly be understood as a kind of desire (to communicate something in a specific way), it may seem that for Goldie ‘interests’ equates desires. However, this understanding of interests does not fit well with Goldie’s overall account of narrative.

Goldie actually points at a different interpretation of interests elsewhere. “Things matter to people, and a narrative involving people can capture the way things matter to them”, he says (Goldie 2012: 23). If narratives reveal what matters to people, then they must reveal more than what people want. Given that maximal understanding goes beyond beliefs and desires, to explain interests as desires does not correspond with the aim of Minimal Narrativism.

It makes more sense to understand interests as the things that matter to people; what people care about. That is what grabs people’s attention when they are giving an explanation about a situation – that is what becomes salient to them. The role of salience becomes increasingly relevant if we really want narratives to have any special connection with understanding the world.

3.2.2. Expanding salience

Earlier, I said that Mrs Jones can give different narratives about her meeting with the consultant. However, when analysing Mrs Jones’s understanding in that example, we are already singling out an event: the meeting. This approach fails to acknowledge that the very fact of singling out specific events is part of the narrative process. Given their interpretive nature the meaningfulness of narrative products – their potential to provide with an explanation – depends partly in their very form. As Mary Jean Walker
says, “what gets counted as “an event” will only be determinable within some interpretive context”, she says (Walker, 2012: 65). I illustrate this with an example.

Let us imagine a person, Ziggy, who is walking across a street at the same moment that a car crashes into a billboard. Applying Lewis’s and Goldie’s account to the example, if Ziggy is prompted to explain the accident he might quote different causes: the tyre, the drunk driver, etcetera. But here, we are already singling out the car crash itself when asking Ziggy. Imagine that Ziggy’s friend, Billy, was also there with him. They both see the crash scene unfold and Ziggy exclaims: ‘Billy, do you see that?!’; and Billy replies: ‘Yes, I can’t believe IKEA has a 50% discount in all stock!’. Although they are both in the situation and they are both witnessing the accident, the accident is not salient at all for Billy. He is, instead, oblivious of the unfolding tragedy, his attention focused on the IKEA billboard above. Ziggy is perplexed: ‘But man, those people in the car!’; to what Billy replies: “Yeah I’ve seen it, but you just have no idea how long I’ve waited to be able to afford a new futon”.

In this case, we have two different interpretations of a situation as two distinct events: Ziggy’s witnessing of a horrible car accident versus Billy’s joyful realisation that he will, at last, be able to afford a new futon. We have two fundamentally different narrative products which bear no comparison with each other, since we cannot ask what about the car accident was salient for Billy: that is not the event he is interpreting. That is what Walker’s statement about what gets counted as an event depending on interpretation means. As it is posited now, Minimal Narrativism cannot explain much about this unusual difference of interpretation, because it is mainly focused on explaining actions and evaluations which have already been singled out. But if narrative theory aims at a maximal understanding of persons, it should not stop at explaining how we can make sense of actions – it should also explain how we can make sense of the very fact that people understand the world in such different ways as Ziggy and Billy do. The difference between Ziggy and Billy is that they find different features of the same situation salient. As it stands salience is underdeveloped in Minimal Narrativism. Goldie’s grounding of salience on interests as mere desires does not correspond with the maximal understanding aim.
3.2.3. The connection between salience and defining traits

To fully explain the notion of salience, I depart from Lewis’s account of causation and look instead at a field where the notion of salience is widely discussed: virtue epistemology. Virtue epistemologists working on explanation have long pointed at salience (‘explanatory salience’ is their term) as one of the features to track knowledge. According to Richard Y. Chappell and Helen Yetter-Chappell,

[salient things are attention-grabbing. A feature of the world is salient to an agent at a time \( t \) in so far as that feature exerts an involuntary draw on the agent’s attention at \( t \).] (Chappell and Yetter-Chappell, 2016: 449, emphasis in original)

There are some general criteria that can be drawn in order to establish what makes certain features of the world salient. John Greco draws from the examples of a house fire and Lewis’s car crash to explain salience:

What factors govern the distribution of explanatory salience? That is not well understood, but we can say some things about it. First, explanatory salience is often governed by relevant interests and purposes. Thus one reason we look to sparks rather than oxygen to explain the warehouse fire is that we more easily control the presence of sparks than the presence of oxygen. Likewise, we readily attribute a crash to the drunk driver, presumably because we have an interest in regulating drunk driving. Second, explanatory salience attaches to the odd or unusual. Thus suppose that a fire starts in an area where there is usually no oxygen, perhaps in some special space in a chemical lab. In that case, we might very well attribute the fire to the presence of oxygen. (Greco, 2012: 7)

There are then two factors which may involuntarily draw someone’s attention when aiming at an explanation: a person’s own interests and an oddity in the situation. Here, I am trying to make a connection between interests and salience, so I will not be discussing oddity any further –neither will I reject or endorse the virtue epistemologists’ claims in that respect.
As we have seen, Goldie recognizes that a person’s interests ground what she finds salient. Chappell and Yetter-Chappell have a richer understanding of interests than mere desires. Salience, they say, is an agent’s disposition to attend, which depends directly from the psychology of that agent (Chappell and Yetter-Chappell, 2016: 450). But the psychological basis of salience entails that salience is not rooted solely in a person’s desires. Instead, “what we find salient reveals what we care about” (Chappell and Yetter-Chappell, 2016: 449).

This connects salience directly with the concept of defining traits that I introduced in §2.1.3. Defining traits are features of a person’s self-concept according to which that person’s life is worth their living. What we have in the example of Billy and Ziggy are two different interpretations of events from two people that are very different from each other. In the framework of this thesis, this means people with very different self-concepts. If Minimal Narrativism aims at maximal understanding, it is an oversight to not take into account that narratives do not provide with an explanation of what people do, but also reveal what people find salient. That is the basis for my argument on the non-trivial character of narrative, which I develop next.

3.3. Self-concept, meaningfulness and salience

In this section, I bring together the roles of salience and meaningfulness in relation to the self-concept. I explain that salience and meaningfulness ground the feedback loop between self-concept and action. I do that by drawing from Walker, whom I quoted earlier as crucially saying a person’s interpretation of the world influences what gets counted as an event.

3.3.1. From the self-concept to action

In a series of papers, Walker develops a narrative account of self-understanding that opens a path for a non-trivial account of the narrative self-concept. Walker (2012) reviews the narrative literature in order to extract claims on self-understanding without committing to the Strong Narrative requirements. Self-understanding, Walker says, involves selection and interpretation. “We remember some events and

---

22 Specifically, she draws from Taylor (1989: 25-52), Schechtman (1996: 96-99), Dennett (1991: 412-430), Velleman (2005) and Lindemann [Nelson] (2001). The fact that many of them are Strong Narrativists does not mean these claims are incompatible with my view, given that they are claims about self-understanding.
forget others, and the way in which we do this relates to what meaning those events have within our continuing story” (Walker, 2012: 65). The way a person connects events in her life is an expression of a person’s character and approach to the world (Walker, 2012: 65). In that way, Walker’s view is very similar to the Minimal Narrativists’. However, Walker offers more details on the claim that our selection and interpretation and events influence the way we see ourselves – our self-concept. (Walker, 2012: 65). According to Walker, the self-concept

encapsulates our reasons for action, since these reasons are related to our understanding of our selves and our pasts. Thus, it frames and guides our decisions about how to act. Our narrative self-conception thus feeds into behavior [sic] … There are thus causal connections in both directions between how one’s life goes, and one’s narrative self-understanding. (Walker, 2012: 65; emphasis added).

What Walker is saying here is that there is a feedback loop between our self-concept and our actions. Who we think we are influences how we see the world (including how we see ourselves). Our individual way of seeing the world (and seeing ourselves) influences what we do. What we do, in turn, influences what we think about the world and ourselves – it feeds into our self-concept. Walker’s claims patently hinge on the role of narrative in self-understanding: ‘the way in which we do this relates to what meaning those events have within our continuing story’; reasons for action are ‘related to our understanding of ourselves and our pasts’.

Walker says that “one’s self-conception, encapsulating one’s various characteristics… directs one towards particular practices and patterns for action” (Walker, 2018: 9). She gives more details on that claim:

[O]ur self-conception shapes our actions and decision–making, as well as how we understand what our choices are. As a result, our actions tend to form patterns over time, to the extent that our self-conceptions retain some similarity over time; and will have coherence with each other, to the extent that our self-conception is internally consistent. (Walker, 2018: 9)
For example, if being a philosopher is one of a person’s defining traits, that directs her actions: she goes to conference dinners, applies to conferences and spends her summer holiday writing twelve drafts of a paper about metaontology. She continues doing this, year after year, so being a philosopher continues being one of her defining traits: her self-conception is internally consistent. If we go back to the previous example of Mrs Jones, I set one of Mrs Jones characteristics as having an inferiority complex. It is because she has an inferiority complex that she repeatedly fails to speak up when she is confronted with an abusive character. This feature of her self-concept not only influences what she does but also the way she interprets what she does and may do. When we read Mrs Jones’s internal perspective, we see that she was scared of saying “something stupid”; that is Mrs Jones’s own interpretation of whatever she could say, informed by her inferiority complex. This limits her choices: speaking up probably does not even feature in her decision-making process when she is trying to decide what to do. Mrs Jones keeps doing this every time she finds herself in a confrontational situation, and this pattern results in her self-conception being consistent over time. In that way, actions contribute to the consistency of the self-concept over time; and the self-concept originates a pattern of actions.

3.3.2 The roles of meaningfulness and salience

Now that I have laid out the foundations of the connection between self-concept and action, I can start developing the role of narrative in drawing this connection. My first aim will be to show that meaningfulness and salience ground the two ends of the feedback loop. To do so, I look at how Walker gives the example of someone who has the defining trait of being very organised:

[S]omeone who views him- or herself as very organized may tend to remember instances that show him or her this way, more easily and more often than instances that show the person to be disorganized. Self-understanding will be selective with regard to what events are included or emphasized. Further, how we remember events is affected by our view of ourselves and by the way we understand those events to relate to other events. For example, an event showing the person to be organized may be connected thematically to other similar events within their narrative—even if the same events could plausibly have a range of other meanings. Thus, our self-narratives are interpretative, and we may
connect events in our lives together in a range of different ways. Since this interpretative activity also leads us to view ourselves in certain ways, it affects how we behave, becoming enacted. If the very organized person approves of this characteristic and interprets it accordingly, it may lead that person to continue attempting to be so. If the person comes to interpret the characteristic differently, say as unnecessary officiousness, perhaps the person would make an effort to alter it. In either case, self-interpretation occurs and plays a role in directing action. (Walker, 2012: 65).

If we look at the process Walker describes, self-reflection looks akin to form-finding, i.e. self-understanding through the creation of narratives. When we engage in self-reflection, we select and interpret different events which unfold over time. This is what Goldie called narrative thinking: we formulate narratives about ourselves, and these narratives have the feature of meaningfulness. They have the potential to explain why we do what we do. Thus, narratives can explain reasons for action through internal meaningfulness. Also, narratives have the potential to explain why we think about what we do in the way we do. That is, they can explain what we think about what we have done — external meaningfulness. What Walker captures that is not explicitly formulated in Goldie’s account on meaningfulness, is that obtaining these explanations affects action. Because of the way we understand our actions – from the events we select and interpret to the kind of explanations we obtain– our self-concept remains internally consistent, which in turn results in a continuation of actions that conform to the initial pattern. So meaningfulness — the ability of narratives to explain— grounds one end of the feedback loop: the direction from action to self-concept. This is how narrative self-understanding has traditionally been understood, and it is the main point of contention from anti-narrativists — given that they reject form-finding is required for self-understanding.

To explain the loop in the direction from self-concept to action, Walker draws from Ricoeur’s classification of actions into ‘higher-order’ and ‘lower-order’ actions. Higher-order actions are at the top of the hierarchy because they contain lower-order, or basic, actions. They are usually long-term goals, or actual patterns of actions, like pursuing a profession or being a believer in a religion. However, Walker warns, people change. Their higher-order goals change (for example, they change
profession or lose their faith), and that entails a change in their actions. Moreover, contingent events also play a part: one might get fired and no longer be in a position to pursue their profession. Personal change happens when there is a break in the feedback loop from self-concept to action. Walker gives the example of a person who experiences a religious conversion:

The conversion constitutes a break across which … the convert’s characteristics, including her values, goals, and reasons for action, alter substantially enough to change what higher-order actions she pursues, and what patterns her actions form (Walker, 2012: 9).

Think, for example of the philosopher above. Imagine she finds herself increasingly thinking that metaontology is a waste of time, dreading conference dinners or leaving a pile of new papers in her field unread. After a while, when engaging in self-reflection she realises that ‘being a philosopher’ is not a defining trait in her self-concept anymore. This realisation, I argue, is enabled by a change in patterns of salience.

Imagine the philosopher, before she becomes disillusioned, is asked by her partner how her day was. The philosopher got up, went to her office to work on a specific metaontological puzzle, had lunch with some colleagues and watched a clip of PMQs. To the possible dismay of her partner, when the partner asks how her day was, the philosopher describes the metaontological puzzle in detail and the different solutions she has been trying to apply, not mentioning any event not related to metaontology. She does not refer, say, to the fact that she was surprised about the sudden coherence in the Prime Minister’s speech she watched on TV, even if at the time she found it surprising. However, someone whose profession is that of a speechwriter would find that event astonishing and particularly salient, probably featuring first in their answer to the question of how their day was. Their attention was involuntarily directed to different things in virtue of their characteristics, and this features in their understanding of the world.

What Walker’s approach does not take into account is that change in self-concept does not necessarily change patterns of actions. The disillusioned philosopher may continue in her post, going to conferences, filling in the REF and looking at metaontological puzzles. What does change is that things related to philosophy do
not become salient anymore when she aims to explain the world (for example, how her day was) and herself (what are her defining traits). Salience grounds the feedback loop in the direction from self-concept to action, since it is through a change in salience that a person can arrive at identifying a break in the consistency of her self-concept (that is, that his actions do not match his self-concept).

Lamarque’s and Strawson’s objections to narrative theory do not identify the role of salience, its connection to the self-concept and from there to action. The objection to narrative theory through an objection to form-finding does not capture the role that narrative has in understanding the world. Narratives do not only allow maximal understanding of people’s reasons for action. Narratives have a role in explaining how people see the world through a loop between the self-concept and action that begins in salience at the selection of interpretation of events and concludes in meaningfulness through the explanation of people’s actions, via a connection with a person’s self-concept and back again. Even if I have not yet offered an argument for the non-triviality of narrative, I think that the fact that anti-narratists do not identify the whole influence of narrative in the process of understanding persons does by itself contribute to a deflection to the triviality challenge. In any case, in the next section I offer several arguments to set the challenge aside altogether.

3.4. Defusing anti-narrativism

A stubborn anti-narrativist could try to go in two directions. Firstly, they could argue that there is no need to create narratives (i.e. no need of form-finding) to explain the feedback loop between self-concept and action. Secondly, they could argue that this cannot be done within a Minimal Narrativist view, given that Minimal Narrativism does not require form-finding and is thus trivial. I will provide with an answer to these next.

3.4.1. Re-interpreting diachronicity through defining traits

In Strawson’s initial formulation of the anti-narrativist challenge, he rejected diachronicity understood as diachronic self-experience: “one naturally figures oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future” (Strawson, 2004: 430). To be an episodic is the opposite of being a diachronic for Strawson, but that does not mean that episodics are necessarily disconnected from their past:
Faced with sceptical Diachronics, who insist that Episodics are (essentially) dysfunctional in the way they relate to their own past, Episodics will reply that the past can be present or alive in the present without being present or alive as the past. The past can be alive—arguably more genuinely alive—in the present simply in so far as it has helped to shape the way one is in the present, just as musicians’ playing can incorporate and body forth their past practice without being mediated by any explicit memory of it. (Strawson, 2004: 432)

In chapter 2, I explained that Minimal Narrativists do not require diachronic self-experience understood in that sense. Strawson says that without that requirement narrative theory is trivial. However, Minimal Narrativism can require diachronicity in a different way. I bring back Walker’s example of “dramatic personal change” (2018:5) where a person undergoes a religious conversion. This is the full articulation of the example:

Take a case in which a person undergoes a sudden religious conversion. This occasions significant alterations to her traits, goals, plans of action, moral orientation, and self-conception. The conversion precipitates changes in the convert’s characteristics, and the relations she takes up toward them. These effect alterations in her approach to the world and how she organises her experience. She re-interprets her past, and her expectations for her future, developing a “new narrative” about herself. Different features of her experience become salient to her, and her behaviour alters in ways that reflect this. As she enacts her new self-understanding, she constitutes herself as a different person in the characterisation sense, from the point of conversion onwards. (Walker 2018:5)

Since I am staying away from any claims on personhood, we can understand the claim of “constitutes herself as a different person” as ‘acquires a different self-concept’, in virtue of the change of her characteristics. Walker’s example of the

---

23 It is not clear whether (Galen) Strawson uses the expression of ‘the past being alive in the present’ in the same way (Peter) Strawson, who is a narrativist (and his father), uses it in a quote I included in chapter 2: “Non-actual perceptions are in a sense represented in, alive in, the present perception” (Strawson 1974: 53, quoted in Goldie, 2012: 119). They do however capture the same idea, and in this section I give a way to understand it from a Minimal Narrativist stance.
conversion is reminiscent of many examples of personal change in the literature. In §2.1 I gave Parfit’s example of the Russian nobleman. Schechtman has another widely discussed example.

Schechtman says that characteristics can be attributed to an agent to different degrees (Schechtman, 1996: 76). In a famous example, Schechtman describes the relation that three women (three ‘matrons’) have with their past selves. The first one, the ‘serious matron’, is a responsible adult woman who looks at her past as a party girl and feels like she is a completely different person: not only she does not behave like her younger self, but in a Strawsonian fashion, it seems like those memories belong to someone else. The second one, the ‘somewhat-less-serious matron’, has also changed her youth habits, but does not feel disconnected from her past: she is able to feel like the party girl when she evokes the memories of those days. The last one, the ‘mortified matron’, is in the same situation as the somewhat-less-serious matron, with an added ingredient: the memories of her younger days “fill her with shame and disgust” (Schechtman, 2001:105). The mortified matron is able to recollect her past memories and recognize them as hers, but she repudiates the characteristics she had then. In these three examples, we see that each present matron has a different attribution relation to the past characteristic of being a party girl. The serious matron does not attribute to herself (her current self-concept) the past characteristic. The somewhat-less-serious matron attributes that characteristic to herself, and approves of it. The mortified matron attributes the characteristic to herself, and disapproves of it.

I will not get into the details of Schechtman’s view regarding the example, since it is formulated within Strong Narrativism. However, here we can draw a parallel between Schechtman’s serious matron and Walker’s convert, given that both people seem to have undergone dramatic personal change. Initially, it seems like the convert has no attribution relation with her past self-conception as a non-believer, like in the case of the serious matron or like the case of the Russian nobleman. They are now ‘different people’; they have radically different self-concepts. However, Walker challenges that intuition:

The convert is somewhat disunified with her pre-conversion self in the sense that her self-conception, patterns of action, and ways of experiencing the world have altered. But she is unified with her pre-
conversion self in the sense that her self-interpretation, in making sense of the changes she has undergone, enables her to understand the pre- and post-conversion self-conceptions in light of each other. (Walker, 2018: 10)

The convert’s new self-concept will motivate a new set of behaviours, beliefs, desires and so on. It will also motivate a re-evaluation of her past self-concept as a non-believer, shedding new light at how she used to understand that part of her life. This re-evaluation will be done through a certain relation (alienation, repudiation, endorsement) which in itself will feature in the organising of both her past and present experiences. She will understand her self-concept, and acts accordingly, in a very different way depending on whether she feels repulsed by having led a non-spiritual life, whether she sees the change as the next natural step in the evolution of her life, or whether she longs for those days when she did not worry about the afterlife. In that way, her past characterisation of a non-believer (and not, say, an unrelated one like having been a party girl or a construction worker) plays a crucial role in her current self-concept. The self-concept may be a new one, but that does not mean that there is necessarily a complete rupture with the past self-concept. Without the past self-concept, the actions, values, beliefs and desires which are in a feedback loop with the present self-concept would be very different.

In this sense, Walker notices an element of personal change which does not usually appear in the discussion. There is no recognition of the fact that if any of the matrons were not looking at the party girl as a matron, but as a CEO of a party-organising company, they would think about the party girl very differently. Having a specific present self-concept –in this case, the matron– influences the way that a past self-concept is interpreted. Having had that specific past self-concept –in this case, the party girl– is crucial for defining the particular way the person relates to those past behaviours and attitudes; and therefore her current self-concept (Walker 2018:7).

The specific attitude which is contained in the relation one bears to a past characteristic is also part of the current characteristics one has. The serious matron’s self-concept is partly composed by her alienation from her past characteristic as a party girl; the same goes for the somewhat-less-serious matron’s endorsement and the mortified matron’s repudiation.
Here we have an answer to what could it mean that ‘the past is alive in the present’, as Strawson claims. There is no need to deny episodic self-experience since, as I have repeatedly stated, self-experience differs from self-understanding. Strawson may very well deny that he experiences himself diachronically, or insist that memories of his past lack any sense of mineness. However, the very fact that he is alienated in this way from his memories will affect the way he understands his current self-conception—apparently, following his own claim, as someone does not care about his past in any special way. That informs his current actions—becoming a staunch opponent to narrative theory. Equally, if the serious matron was not alienated from the party girl (i.e. if her memories of the party girl felt like her own) she would have a different self-conception. She would actually be a somewhat-less-serious or a mortified matron.

Given that self-concepts are embedded in a feedback loop with action, the fact that they cannot be understood without reference to past self-concepts (that is, in a diachronic way) means that the feedback loop—the actual connections between self-concept and action—can only be best explained in diachronic form. That is, in the form of a narrative, which let us not forget is not a mere causal account, since narratives are a route for maximal understanding which is not accessible through merely causal accounts. Creating narrative products is the best way to understand these connections, so the role of narrative in understanding persons is not trivial. Next, I put forward a second argument in favour of the non-triviality of the role of narrative in understanding persons.

3.4.2. Re-interpreting story-telling through narrative models

In the previous two chapters, I have repeatedly endorsed Helm’s criticism of the ‘individualist conception of persons’. People do not exist in isolation: they exist in interaction with other people and within society. Helm is mainly concerned about how interacting with other people in close relationships shapes a person’s identity. What he overlooks is that the social concepts that people share also contribute to shaping a person’s identity—in my discussion, a person’s self-concept.

I said in §2.3.3 that the Strong Narrativist requirement for story-telling should be re-interpreted. Within Minimal Narrativism, it is not a requirement to live one’s life as a quest (as some Strong Narrativists suggest), but the acceptance of the fact that
certain stories influence the way we understand the world. Later, in §3.1.2 I showed that Hutto calls these stories ‘narrative models’. Now that the feedback loop between the self-concept and action has been established, I am in a position to properly assess the relevance of narrative models in a non-trivial narrative view.

Hilde Lindemann is plausibly the narrative theorist who has developed the idea of narrative models the most. She calls them ‘master narratives’. In order to explain the role that they play in forming the self-concept, let us look at this example she gives about pregnancy:

[I]dentities—perhaps especially the pregnant woman’s identity—set up socially shared normative expectations. Recall that our identities serve as guides for what we are supposed to do, and we treat ourselves and others according to our narrative understanding of who we and they are. Many of the narratives that constitute the identities of pregnant women are personal and particular… More of the stories, though, are the master narratives, ubiquitous in movies, TV, magazines, and other media, that show pregnant women undertaking what in the United States has become an astonishingly large number of material and mental practices: announcing the pregnancy to family and friends, watching what they eat and drink, posting status updates on Facebook, putting themselves under a doctor’s care, buying or borrowing baby furniture and clothing, reading What to Expect When You’re Expecting, talking about their pregnancies and plans for delivery to sometimes even the most casual acquaintances, putting wish lists on the Target and Babies “R” Us gift registries, monitoring their blood pressure regularly, and on and on and on. (Lindemann, 2014: 49)

There are certain practices—like pregnancy—that come with a social script of sorts; a series of “material and mental practices” that not only are social expectations, but that become reasons for action. Obviously, many of these social scripts include reasons for action that are independently justifiable. A pregnant woman has reasons to put herself into doctor care not only because that action is given to her by script, but because she has independent reasons to promote her and the foetus’s welfare. She also has reasons to tell her friends about her pregnancy not only because it is in the social script, but because it is part of their relationship to share important events
in their lives. But regardless of independent reasons, once a woman gets pregnant, she will have a narrative model of pregnancy which become a possible, and often generalised, route of action.

The relevance of narrative models is even more patent in the case of people in situations of oppression. Many racialized people in countries with a white majority are understood, in the minds of many people, within narratives of violence and crime—a narrative model of ‘what a racialized person does’. These narratives irremediably influence the way they are treated in interpersonal interactions:

Oppressive master narratives that enter into social group identities commonly make it impossible for the people bearing those identities to express themselves adequately, and then, of course, what they say and do won’t get the right kind of response. It’s not that people in the stigmatized group can’t utter the words or perform the actions that reveal their mental states, but that their oppressors refuse to acknowledge their right to have those states. The woman’s No translates as playing hard to get, is taken as meaningless noise, or enhances the sense of conquest… The black teenager’s hoodie gets registered as gang membership and the self-appointed neighborhood [sic] watchman shoots to kill. (Lindemann, 2014: 115)

Understanding (not justifying) the neighbourhood watchman’s action partly requires an investigation into the racist narrative model fed into that action. Understanding the pregnant woman’s reading What to Expect When You are Expecting partly requires an investigation into the narrative model for pregnancy among Western middle classes. This is because these courses of action become salient partly in virtue of them being social practices fixed in narrative models.

Again, anti-narrativists may oppose that they live their lives according to a genre, but they can hardly deny that social expectations shape their lives. Given that these become part of the feedback loop between the self-concept and action, rejecting story-telling (and thus narrative theory) is rejecting a route to maximal understanding. Identifying narrative models is the best route to define how these social expectations translate to a person’s self-concept, which stands in a feedback loop with a person’s actions. Thus, this is another reason why Minimal Narrativism is not trivial. Finally,
I make some clarifications on the non-trivial narrative view on the self-concept I have drawn here.

### 3.5. Final observations on truth and revision

Strawson, as seen in §2.3, identifies the requirement for revision as one of his objections to narrative theory. According to him, the requirement of revision is that “one will also have a tendency to engage unconsciously in invention, fiction of some sort – falsification, confabulation, revisionism – when it comes to one’s apprehension of one’s own life” (2004: 443). I already said that no narrativist view makes falsification or confabulation a requirement. However, the possibility of truth in narrative is a very prominent topic in the literature, usually presented as an objection to narrative theory. People tell stories about themselves and others that are not true; by putting events into narrative form we sometimes make up or directly falsify the truth. Although I am not going to offer a fully-fledged argument on the possibility of truth in narrative, there are two truth-related features of narratives that I think are important.

Firstly, the danger of narratives could be what Goldie calls dangerous fictionalisation theories (2012: 151). That is, the idea that our lives should look like fictional narratives. Like Goldie says, these are dangers that should be recognized, but there is nothing in narrative thinking which makes fictionalising necessary. Understanding persons through narratives, he admits, is “messy and imprecise” – but the reply is that “this is just what it should be, given that life itself is messy” (2012:173). Given also that the topic of this thesis, romantic love, is maybe among the messiest of the messes that feature in a person’s life, I endorse Goldie’s view on imprecision and move on from this objection without further discussion.

Secondly, narrative seems to be prone to confabulation, understood as the creation of false memories or false interpretation. Here, one possible answer is that confabulation can be beneficial. In a series of papers, Lisa Bortolotti has drawn from empirical evidence to argue that confabulation may actually bring epistemic benefits, that is, it may increase the chances of achieving true knowledge (Bortolotti 2018;

---

24 Goldie identifies four fictionalizing theories: “we see ourselves as plotting our lives; we find agency in the world where there is not; we seek narrative closure where it cannot be found; and we transport notions of genre and character from fiction into real life” (2012: 171-172).
Puddifoot and Bortolotti 2019). It is not necessary to go into the details of that idea, however, because my line of response is simpler than that: there is nothing about narrative accounts that makes them more prone to confabulation than causal accounts. Larry can be confabulating about his past in the same way that Perry or Harry can. This is not a theory of how we achieve true knowledge about persons, but about how we understand persons. Again, I relate this to the general topic of the thesis. This is not a thesis about knowing whether we are in love. Whether and how we can know for certain if we are in love with someone is definitely an interesting question, but is not the one I am concerned with.

More importantly, there is a feature of revision that the anti-narrativists overlook. We can have narratives about ourselves which we revise in the light of other people’s observations. The revisable nature of narratives seems like an immediate disadvantage for self-knowledge. In the specific case of romantic partnerships, we can have a narrative about our relationship being one of bliss and, when someone points out that we are actually in an abusive relationship, become aware of the truth. This is, indeed, a risk. But what the anti-narrativists do not realise is that revision is precisely one of the advantages of narrative theory. Steve Matthews and Jeannette Kennett have identified that the problem of false narratives is especially poignant in intimate relationships, given that these relationships partially consist in the sharing of our respective stories (Matthews and Kennett, 2012:301). The reason why lying is particularly damaging in intimate relationships is that engaging in a relationship means, precisely, to open our self-concept to re-interpretation by our close ones (Matthews and Kennett, 2012: 312). Kevin Harrelson explains it well in his defence of the role of revision in self-knowledge (he has a specific example about a friend who is becoming suspicious of his brother, who has falsely adopted the narrative of a ‘self-made man’):

Our self-concepts motivate biased, perhaps even misleading, recollections of the past. Nonetheless, it follows from this only that narratives based on memory serve as evidence in self-knowledge only if other people broadly corroborate what we tell—the other people not being subject to the same selection bias in regard to our case. We attain self-knowledge mainly by sharing stories with others and allowing the stories to be corrected by them. Hence my friend can remind his brother
that the three generations of lawyers in their family have contributed to his academic success, and your therapist (or your friends) can assist you in uncovering episodes that your self-narratives exclude. (Harrelson, 2016: 172).

Revision can then go both ways: yes, it can corrupt self-understanding if we create false narratives about ourselves; but given that we tend to have a confirmation bias towards the beliefs we have about ourselves, it can also correct false narratives. Revision is then a double-edged sword, both a problem and the solution to the problem. But more importantly, Harrelson reveals that revision, in the case of romantic love, is a necessary condition:

Meeting a stranger… became meeting my best friend after a subsequent sequence of events was completed. And a consequence of this is that stories are as revisionary as Strawson feared, except that the fact does not pose a problem for their epistemic value. (Harrelson, 2016: 176)

When we make a friend, when we are in love with someone, those are not things that happen overnight. They happen over time, through a ‘sequence of events’, which at the time we often fail to identify as part of the process of befriending or falling in love with someone. So revision is not only a potential solution to a problem, and not only a requirement in personal relationships like Matthews and Kennett say, but a requirement for being in love. My account on being in love, which I develop in the next two chapters, rests on these claims made by Harrelson, and Matthews and Kennett. For now, I think I have sufficiently shown that the risks that truth pose to narrative are then not that great, and the fear of revision, unjustified. With that, I conclude the part of the thesis focused on finding a plausible narrative theory of the self-concept. In the next two chapters, I develop an account on being in love which rests on this theory.

Conclusion

Throughout a detailed analysis of the strengths and weaknesses of Hutto’s and Goldie’s Minimal Narrativism, I have arrived at a view on the narrative self-concept which avoids the triviality challenge. From Hutto and Goldie, I have accepted that narratives are a route to maximal understanding. Given that they reveal the
evaluative and the emotional import that events have for the person narrating (and, further ahead, the narrative models influencing them), narratives have a higher explanatory power of reasons for action than belief/desire–based accounts or causal accounts. I have pointed out that in order to fully achieve the role Minimal Narrativism gives to narrative, we need to incorporate salience to the account. Salience grounds a person’s selection and interpretations of events, and is itself grounded in a person’s characteristics.

I have argued that the self-concept is in a feedback loop with action that is grounded in both ends by meaningfulness and salience. Meaningfulness grounds the loop in the direction from action to self-concept: a person’s understanding of her actions influences her self-concept. Salience grounds the loop in the direction from self-concept to action: a person’s selection and interpretation of events influence her actions. This feedback loop is diachronic without the need for self-experience to be diachronic: current characteristics are partly constituted by past characteristics. In that sense, self-concepts are diachronic in a way that avoids Strawson’s objection. Given all these features of the self-concept, and together with the influence of narrative models in action, narrative theory is not trivial. The self-concept, which is narrative, articulates a person’s understanding of herself, the world and others. I now move on to show how this account is illuminating in order to explain what being in love is.
Chapter 4

Love as mutual shaping

Now that I have offered a plausible account of the narrative self-concept, I move on to show that the self-concept and narrative theory are both crucial to describe what being in love is. In this chapter, I show how the right intuitions of the caring view can be incorporated into a view on love which is based on the lover’s self-concept, avoiding the challenges I presented in chapter 1.

In §4.1 I discuss Dean Cocking and Jeanette Kennett’s (1998) ‘drawing view’ of friendship, according to which it is a necessary condition for friendship that one’s choices are directed by a friend, and that one’s self-concept is shaped by a friend’s interpretation. Helm makes several objections to the view, which are overcome by bringing in Amélie Rorty’s (1987, 2016) and Alexander Nehamas’s (2010) compatible views. I call the view resulting from these combined accounts ‘mutual shaping’. However, mutual shaping still faces an objection: it is not clear why one should accept direction and interpretation from a loved person any more than from anyone else.

In order to overcome this objection, I incorporate Karen Jones’s (2008) view on being in love as a trajectory. According to Jones, being in love is a trajectory-dependent property: the fact that a person sees another as a loved person at \( t \) depends on attitudes and events at \( t-1 \) and/or \( t+1 \). Trajectories have a narrative shape then, given that events make sense in light of each other. In this view, the lover is more likely to make choices that continue the love trajectory. By merging Jones’s claim with the mutual shaping view, one accepts direction and interpretation because one is in love.

In §4.3, I present my objections to the mutual shaping view. I show that in cases of conflict, love seems to justify itself. This problem is overcome if being in love is understood as a characteristic of the narrative self-concept—a view that unlike mutual shaping, acknowledges that love not only shapes one’s self-understanding, but also one’s understanding of the world.

A disclaimer: some of the arguments in this chapter are not exclusive to being in love. The mutual shaping view is mostly developed in the context of friendship, not
romantic love. However, everything I discuss in this chapter is to be understood as applying to romantic love too, regardless of the subject matter discussed by the authors I bring in. My main aim is here is to show how a view based on mutual shaping in the context of the self-concept avoids the problems of the caring view. In chapter 5 I describe how this view can apply to romantic love specifically. For that reason, and because the dyadic ‘friend/friend’ is due to cause confusion, I do use the terms ‘lover/loved person’ and ‘lovers’ interchangeably with ‘friends’.

4.1. Mutual shaping versus caring

Let us remember the problems of the caring view, which I discussed in chapter 1. If love requires the incorporation of the loved person’s interests to the lover’s own identity –like Frankfurt argues–, then we cannot count as love cases where the loved person’s interests do not become the lover’s own interests. Helm proposes a caring view where the loved person’s interests do not become the lover’s own interests. Instead, love requires a pattern of emotions towards the loved person’s values. However, as I showed, that view cannot accommodate love in spite of problematic identities, love in spite of opposing identities, cases where the lover does not care about the lover’s values, or cases of conflict. On that basis, I proposed a change of direction: making love about the lover.

The mutual shaping view does precisely this, incorporating what the caring view gets right while avoiding the caring view’s problems. Briefly, the mutual shaping view defines love as a change in the lover’s self-concept which happens relationally, through the relationship with the loved person. It is then compatible with the fact that some of the loved person’s interests become the lover’s interests and the fact that self-concepts are formed relationally. It is also, as I will show, compatible with a conception of well-being that does not conform to universal or moral standards.

4.1.1. The basics of mutual shaping

Dean Cocking and Jeanette Kennett (1998) develop the ‘drawing view’ of friendship, which is based on the mutual shaping of the self-concept. According to Cocking and Kennett, “as a close friend of another, one is characteristically and distinctively
receptive to being directed and interpreted and so in these ways drawn by the other” (1998: 503). Let us look closely at the concepts of direction and interpretation.

To be directed by someone is to have one’s own choices shaped by that person. It is not a matter of being told what to choose (what to think or what to do), but a matter of our thoughts and actions changing in virtue of the interests of the people we love. In friendship, Cocking and Kennett say, “each person is receptive to developing interests or activities, which they do not already pursue, primarily because they are the interests and activities of the other” (1998: 503–504). For example, one may join a friend to see ballet if they are interested in ballet, even if one is not personally interested in ballet. The choice of going to the ballet is motivated by one’s friend’s interests in virtue of them being interests of a friend, and not of a random person – or a person with whom one may have a relationship that is not as intimate as friendship, like “a new colleague or an elderly aunt” (1998: 504).

By having one’s choices directed by a friend, one’s self-concept may change; for example, ballet may become part of one’s own set of interests. That means that becoming friends with someone can change the lover:

On the basis of this receptivity to my friend’s interests, aspects of my character may change in ways that they otherwise might not have and such changes may persist beyond the friendship. I might get infected by Iris’s enthusiasm for ballet; typically, I will at least be interested in understanding and appreciating it, simply because she loves it. Thus I may change from someone who had a marked distaste for high culture to someone with a genuine appreciation of the athleticism and grace of dance. (Cocking and Kennett, 1998: 504)

The change brought by friendship in the lover is even more patent through interpretation—the other component of the drawing view. When we are friends with someone, Cocking and Kennett point out, we recognise features of their actions and character and make observations about them. We do not need to be similar to someone to interpret their actions and character (Cocking and Kennett, 1998: 509); that explains why sometimes friends are very different from each other. Having our behaviour interpreted by another can result in a change of our self-concept. Cocking and Kennett describe this change in the following example:
So, for example, Judy teasingly points out to John how he always likes to be right. John has never noticed this about himself; however, now that Judy has pointed it out to him he recognizes and accepts that this is indeed a feature of his character. Seeing himself through Judy’s eyes changes his view on himself. But beyond making salient an existing trait of character, the close friend’s interpretation of the character trait or foible can have an impact on how that trait continues to be realized. Within the friendship John’s liking to be right may become a running joke which structures how the friends relate to each other. John continues to insist that he is right; however, his insistences are now for the most part treated lightheartedly and take on a self-consciously ironic tone. And John may be led by Judy’s recognition and interpretation of his foibles to more generally take himself less seriously. Thus, John’s character and his self-conception are also, in part, drawn, or shaped, by his friend’s interpretations of him. (Cocking and Kennett, 1998: 505)

Cocking and Kennett claim that drawing—understood as direction and interpretation—is both a condition and a component of friendship (1998: 506). Firstly, people who are not open to being directed and interpreted cannot have close friends, since they have “a narrow self unable to engage with the world and others” (Cocking and Kennett, 1998: 520). It is then a necessary feature of friendship that the friends are open to engaging with each other in the terms they describe—by allowing themselves to be directed and interpreted. Otherwise, there is no friendship:

It just cannot be the case, then, that the person’s identity and reasons for action where the friend is concerned are altogether undetermined by the friend’s direction or the person’s interpretation of their friend. (Cocking and Kennett, 1998: 524).

However, the necessary condition here is to be open or receptive to be drawn. There is, unlike in the caring view, no volitional or rational requirement to share the loved person’s interests (which are Frankfurt’s and Helm’s caring views, respectively). As I showed in chapter 1, caring views do not accommodate cases where the loved person’s interests are rooted in a problematic identity (like Pete Doherty and drugs); where the loved person’s interests are opposing to one’s own interests (like Lucy
Mangan and her husband’s Tory ideology); or where one simply does not care about the loved person’s interests (like Silvio with Angela’s supporting of Manchester United). Cocking and Kennett’s view does accommodate these cases, given that they give up the understanding of ‘caring for the other’s sake’ as the articulating principle to define love. In the drawing view, there is no ‘for the other’s sake’ requirement, where ‘the other’s sake’ is understood as the other’s identity as a whole:

Being close friends with another does not mean that you will be disposed to be directed by their interests in a wholesale fashion. Apart from the possibility of the friend having immoral, idiotic, or dangerous interests that you would not be prepared to share, you might also just not be prepared, for instance, to go to the ballet with [your friend]. Indeed, it is surely a part of most, if not all, friendships, that each party has some interests that the other will simply never have an interest in sharing. It would be an unrealistic, romanticized view to expect otherwise. (Cocking and Kennett, 1998: 524–525).

By dropping the requirement to love ‘in a wholesome fashion’ – i.e. of love being directed as the whole identity of the lover– Cocking and Kennet are then able to accommodate non-idealised love. Someone can love Pete Doherty in spite of his valuing drugs, with drugs not requiring any pattern of emotions from the lover. However, Doherty also plausibly enjoys going to music festivals or clubbing, and what the mutual shaping view says about direction is that Doherty’s lover will participate in some of these interests. Equally, someone who loves Doherty must be open to his interpretations of her: if he tells his lover that she is kind, or too talkative, or that she spends too much money, all she needs is to be receptive to this interpretation. The same goes for opposing identities: Mangan need not incorporate her husband’s interest in the Tory party, and Silvio need not incorporate Angela’s interest in Manchester United. As long as they participate in other interests of their respective lovers and are open to being interpreted by them, Mangan and Silvio can still be in love.

I have pointed out previously (§2.1.3) that the caring view does uniquely account for some features of being in love which I believe are correct. One of them was the fact that identity is formed relationally and not individually, and that the people who we are in personal relationships with play a crucial role in the shaping of our identity. In
that sense, Cocking and Kennet’s view stands on equal grounds with Helm’s view – the relational nature of identity is one of the reasons they mention in favour of their view (Cocking and Kennett, 1998: 505). Helm does discuss Cocking and Kennett’s account but rejects it in favour of his caring view, on the basis of two objections. Firstly, he argues that Cocking and Kennett give too much of a passive role to the lover. Secondly, he argues that Cocking and Kennett’s view entails an acceptance of the loved person’s damaging interests, which is a failure of concern for the loved person. I discuss these objections next. In order to show the solution, I draw not only from Cocking and Kennett, but also from two other accounts which have many points in common with theirs and offer more explicit solutions to Helm’s challenges – namely Amélie Rorty’s (1987, 2016) and Alexander Nehamas’s (2010, 2016).25 Given that the discussion incorporates views other than Cocking and Kennett’s drawing view, from now on I call the view I have described ‘mutual shaping’.

4.1.2. Change and dynamism

Helm criticises Cocking and Kennett on the basis of the role they give to the lover, which he believes is too passive:

[I]t is unclear what your role is in being thus directed and interpreted by your friend. Is it a matter of merely passively accepting the direction and interpretation? This is suggested by Cocking and Kennett’s use of the word, ‘receptivity’, and by their apparent understanding of this receptivity in dispositional terms. (Helm, 2010: 259).

Cocking and Kennett are, indeed, not very explicit on the role of the lover. This is partly justified by their aim, which is to argue for the active role of the loved person which, in their view, has been mostly ignored (Cocking and Kennett, 1998: 514).26 Helm’s understanding of the view is not the most charitable one. Even if they do not say it explicitly, what the lover does is crucial for Cocking and Kennett, given that

25 Actually, Helm sees Cocking and Kennett’s view as an inheritor of Rorty’s (Helm, 2010a: 259 fn.11); Kennett herself sees Nehamas’s account as a continuation of the drawing view (Kennett, 2017: 275).

26 One of the views Cocking and Kennett target is the Aristotelian mirror view, according to which friendship is based on the mutual recognition of virtue, so the friend becomes ‘another self’. Aristotle is probably the genesis of the ‘caring for the other’s sake’ tenet of the philosophy of love: “Those who wish good things to their friends for the sake of the latter are friends most of all, because they do so because of their friends themselves, and not coincidentally” (NE 1156b9–11).
being directed by the loved person results in the lover doing things—the things that are interests of the loved person. The actions of the lover matter for Cocking and Kennett, so it is hard to defend that the lover is not active. Helm does not either acknowledge that Cocking and Kennett require that the lover accepts the loved person’s interpretations as being accurate. This condition is revealed in Cocking and Kennett’s discussion of how change in the lovers can influence friendship.

According to Cocking and Kennett, a disruption in the receptivity to be directed and interpreted by a friend is a “terminating condition of friendship” (Cocking and Kennett, 1998: 521). With that, Cocking and Kennett depart from much of the rationality approach to love. As I mentioned in the Introduction and in §2.2.1, a great number of philosophers of love are focused on finding the reasons that ground love. Some of them argue that love is grounded in the properties of the beloved; one of the objections to property-based views is the fact that people change over time. The person we fall in love with today may be the soul of the party and become a bore in a few years, but we may nevertheless continue loving this person. Cocking and Kennett say that it does not matter how much the loved person may change: friendship can survive change of its members. What it cannot survive is a disruption in direction and interpretation. If I become impervious to the direction and interpretation of someone (her interests no longer influence my actions; I am impervious to her observations) then I do not love that person anymore. But crucially, I myself have to assess the interpretation of the loved person, which gives me an active role:

At least one of us has changed, it is true, but it is not the fact or even the direction of change that is significant here. Rather, it is that I feel that [my friend] no longer knows who I am. She is not interested in me; her interpretations are either insufficiently responsive to the changes in me, or are sensitive, not to me, but to her political or religious beliefs or to suggestions from third parties. Similarly, my revised interpretations of the other might present her as someone who is beyond any influence from me, or from whom I am no longer willing to accept interpretation and direction. Thus, she and I can no longer be considered companion friends, though it need not be the case that we are now less similar than we were. (Cocking and Kennett, 1998: 522; emphasis in original)
The lover is not a mere passive receiver then of the direction and interpretation. The fact that this gives an active role to the lover is shown more clearly in Rorty’s (1987) account. Rorty defines love as “dynamically permeable” (1987: 402). Permeability is similar to Cocking and Kennett’s drawing: it means that the lovers mutually shape each other’s identities (Rorty, 1987: 402). Dynamism, for Rorty, is more than a response of the lovers to the change in each other’s properties (for example going from being the soul of the party to being a bore). It is a response of the lovers to the changes caused by love: “every change generates new changes, both in the lover and the interactions with the friend” (Rorty, 1987: 402). Rorty gives the example of a couple, Ella and Abe, who are trying to discern where to move in virtue of each other’s careers:

If they moved to New York so Ella could sing at the Met, her love of music would become focused on Verdi or Wagner rather than folk or rap. If they had stayed in Boston to enable Abe to influence NPR music programming, Ella would be singing Gregorian chants in the choir at the Cathedral of the Holy Cross. The very quality of her voice and her pleasure in singing would change. (Rorty, 2016: 351)

This is something that is not explicitly discussed by Cocking and Kennett, which is relevant to understand the active role of the lover in the mutual shaping view. Rorty captures the idea that in virtue of directing each other, their actions will determine other actions they take. And, crucially, sometimes they will be directing each other in opposite directions. In the example, Ella advocates for a move to New York; Abe for a move to Boston. As I said in §1.4, Rorty is one of the few authors that explicitly discusses the fact that conflict is present in non-harmonic love. I gave the example of Lydia and Mema: Lydia is a supporter of Manchester United, which Mema hates. In the mutual shaping view, this is not a factor against love—as long as they are open to being directed by some of each other’s interests. However, it will require negotiation: I said that, for example, Lydia and Mema would agree to go hiking in alternate weeks so Lydia can go to the match every other week. Conflict needs to be actively solved by both lovers. And, given that love itself brings about changes in the lovers, it would require an active adaptation to these changes in order to negotiate these conflicts:
Abe and Ella are active and interactive in attempting to construct as well as determine in the details of their happiness. They do not just discover and adhere to what love demands of them as if they were following a pattern; they themselves develop the expressive details of their love, sometimes by a process of unexpected improvisation (Rorty, 2016: 351).

Hence, Helm’s objection on the passive role of the lover does not hold. First of all, the lover needs to assess the loved person’s interpretation, and accept it as an interpretation that is responsive to the changes she has gone through (as Cocking and Kennett argue). Secondly, given that direction and interpretation entails further change, and given that the lovers may be drawing each other in different directions, the lover needs to actively respond to these further changes and conflicts. The lover is not passive in the mutual shaping view.

4.1.3. Dangerous friendships

Helm posits a second objection to Cocking and Kennett, specifically in the case of love for people with problematic identities. He gives the example of a loved person who values gambling:

If you understand gambling to be wrong, then you fail in your concern for your friend if you blindly go along with him. (Helm, 2010: 259).

As I showed in chapter 1, Helm’s position on love for people with problematic identities stands in contradiction with his view on well-being as the fostering of one’s conception of a life worth living, which is not necessarily defined through moral or universal standards of well-being. The above statement stands on that contradiction. However, it could be objected that I have accepted that conception of well-being too fast in chapter 1. I said that given that we can love people whose values do not conform to universal or moral standards, a theory of love which wants to preserve

---

27 Benjamin Bagley (2015) calls love a process of ‘deep improvisation’. Comparing love with the performing of jazz music, Bagley describes the interactions and the mutual self-creation of the lovers in a way that mirrors Cocking and Kennett’s and Rorty’s view. However, his view is inseparable from universal or moral standards of well-being: “It’s important here that a partner in deep improvisation necessarily values your living in the way you find fundamentally worthwhile and that this constitutes a central dimension of autonomous flourishing” (Bagley, 2015: 504). In chapter 1 I endorsed Helm’s conception of well-being as not conforming to universal or moral standards, which I back further in this chapter (§3.1.3). Hence, I do not engage with Bagley’s view.
the intuition that love is part of a life worth living needs to rest on a compatible conception of well-being. Now, with the mutual shaping in hand, I can offer further support for my position—and reject Helm’s objection.

In *Friendship and Moral Danger* (2010), Cocking and Kennett target their drawing view towards an idea which is widely shared in philosophical circles: the conception of friendship as a route to moral development. Briefly, the idea is that friendship is only true friendship insofar that it contributes to a morally good life. This conception tracks back to Aristotle, for whom perfect *philia* or true friendship consists of mutual goodwill and mutual appreciation of virtue. By making friendship a virtuous activity, Aristotle establishes love as a crucial component of *eudaimonia* or the morally good life—arguably making of friendship the ultimate moral activity (NE VIII/IX; EE VII). The contribution of friendship to the moral life is the foundation for numerous views on love (Cocking and Kennett focus on Sherman, 1993; Thomas, 1989; Friedman, 1993; and Blum, 1993—see also Murdoch, 1970; Solomon, 1988; Velleman, 1999). Cocking and Kennett call out these views for being “highly moralized” (2010: 280). These views are not only misguided: they exemplify a “striking” gap between philosophical accounts of friendship and everyday experiences of the phenomenon (Cocking and Kennett, 2010: 279).

Cocking and Kennett do not deny that friendship can lead to moral development. As seen in §3.1.1, direction and interpretation entail that the loved person’s interests shape the lover’s interests, and ultimately, her self-concept. It follows then that a person of immoral tendencies can befriend a highly morally committed person and change for the best. But what is a possibility in friendship cannot be taken as a condition for friendship. In reality, Cocking and Kennett argue, friends can lead each other to commit morally wrong actions, and their friendship can still be true friendship—and a valuable friendship— if they do. Cocking and Kennett give two reasons for this. Firstly, sometimes friends do immoral things for their friends, precisely out of friendship:

---

28 I say friendship is “arguably” the route to the most moral life for Aristotle because it is a common source of disagreement among Aristotelian scholars whether social life and contemplation are mutually exclusive for the *eudaimon*, on the basis of the tension between books VII and IX with book X of the *Nicomachean Ethics*. See Cooper (1977) for a well-established analysis of Aristotle’s view on friendship. See also Annas (1993) for a rejection of the requisite of virtue in friendship from within Aristotelian scholarship.
It might not be morally praiseworthy of me to spin a tale to my colleague and break my promise to her so that my friend and I can go off to the movies together. But it is hardly the case that I am not being a good friend here. I might be a perfectly good friend. I might just not be a perfectly moral one. (Cocking and Kenett, 2010: 287)

This does not mean that friendship warrants just about any immoral act. When the moral stakes are high, Cocking and Kennett note, most of us would not do just about anything for our friends, like for example murdering someone (2010: 294–295).

Secondly, what attracts us to our friends (i.e. the characteristics that become salient to us and we therefore interpret) are not always morally good characteristics:

Recklessness is not morally admirable, but it might be what I like about you and it may well structure the ways in which we relate to each other and the activities we share. I am just as likely to be directed by your interest in gambling at the casino as by your interest in ballet. (Cocking and Kennett, 2010: 286)

Not only that, but one might change for the worse, morally speaking, precisely as a result of a friend’s interpretation. It might be, for example, that I am only becoming faintly interested in gambling, and when my friend interprets that as “exciting, wild, cool, or hilarious” (Cocking and Kennett, 2010: 286), my self-concept and subsequent reasons for action may end up cementing that tendency. I may indeed gamble more frequently until ‘being a gambler’ becomes a trait of my self-concept, which feeds into my actions—more gambling— and back into the self-concept through the feedback loop. Had my friend not cheered my incipient gambling, I might have just lost interest and moved on to another, less morally reprehensible hobby. Indeed, the potential danger of friendship for character is too common a fear for the parents of teenagers who start hanging out with ‘the wrong kind of people’, to use a common figure of speech (Cocking and Kennett advance this example in 1998: 514). The example might prompt the objection that, like teenagers, when we change for the worse because someone interprets our incipient vices as “cool”, all that is happening is that we are looking for adolescent, irrational validation—it has nothing to do with true friendship. However, that would underestimate the value of
friendship for its members, even friendship that disconnects them from moral and universal standards of well-being.

To illustrate that value, Nehamas (2010) discusses the film *Thelma and Louise*, a quintessential story of female friendship. Briefly, the plot focuses on the journey of the two women in the title. Louise is a submissive wife with a boring life who goes on a short holiday with her friend, free-spirited Thelma who seems to have her life under control. During a stop in their travels, a man tries to rape Thelma. Louise arrives and hesitantly threatens him with a gun. Thelma makes it out of the man’s grip but spitefully says that he wishes he had raped her, so Louise shoots and kills him. That kick starts a runaway from the police which gets them to rob a store, kidnap a policeman and blow up a truck, among other clearly immoral and non-eudaimonistic actions. In the end, cornered by police at the rim of the Grand Canyon, they purposely plunge to their deaths while holding hands, choosing to go on their own terms instead of giving up their freedom.

Nehamas notes how both women change deeply during the journey: Louise loses her insecurity and Thelma becomes kinder. That, and not the moral character of the actions they perform together, is the most relevant feature of their friendship:

> Thelma and Louise are not just friends who do bad things, as friends often do. The friendship they acknowledge to each other before driving into the canyon is a friendship through which both women, especially Thelma, become more admirable because of the bad things they do. It is through bad things that they gradually become equal to each other and to the world and find a way out of the dead ends to which their lives would have inevitably led otherwise. (Nehamas, 2010: 276)

So it not only may be the case that I can have a true friendship with someone who has an immoral character and/or with whom I do immoral things. What Nehamas highlights is that the very fact that we do immoral things together may turn me into a better person. It was probably Thelma’s self-confidence that gave Louise the strength to go to defend her in the parking lot, and it was probably that glimpse of self-assurance that made her shoot the offender. Seeing herself standing up for herself in that way probably triggered a change that was fostered by her friendship with Thelma; we see it realised at the end. There is no need to defend murder here:
it is easy to imagine Louise getting to the same place by doing something less definitive, but also immoral, like shooting the offender in the foot. The important point is that she would have never pulled that gun, and subsequently changed, without Thelma’s influence; and she would have probably never changed if she had not pulled that gun in the first place and gone on the run with Thelma.

As Kennett notes, this is an “uncomfortable” idea (2017: 276) which nonetheless shows how responsiveness to one’s friend is constitutive of true friendship regardless of morality. Since mine is not an argument about morality, I do not need to go as far as endorsing that committing immoral acts with friends can contribute to your moral development—as Nehamas does. All I need is to acknowledge that immoral activities can be the basis for mutual shaping as much as moral activities can be. The mobsters who force their bond through a series of assaults and robberies need not have a less true friendship than the volunteers who become friends in a homeless shelter. The mobsters can potentially interpret and direct each other in the same way the volunteers can.

I must insist: nothing in the discussion above rules out that love can lead to moral development. It may even be the case that moral development happens at its best in love, between family members, friends and partners. I am certainly sympathetic to that idea. Also, defending that friendship can exist on the grounds of immoral action does not mean that such friendship is of the best kind. As Neera K. Badhwar says, Thelma and Louise’s friendship “would have been a far greater good had they had better judgment and not created situations in which doing the right thing by each other required doing the wrong thing by others” (2016, n.p.). The risk of being led morally astray by a friend, Badhwar continues, “depends to a large extent on the character of the friends in question” (2016, n.p). I agree with her. But one can endorse Badhwar’s points while at the same time arguing that it is a mistake to believe that the potential for moral development is the metaphysical key to distinguish between true and fake friendship—if anything like ‘fake friendship’ exists at all. True friendship can happen between immoral people. Maybe their friendship is objectively

---

29 Even if this were true, it would not be exclusive to friendship: think of the redemption tale of the former addict who becomes a pillar for her community as a result of her history of addiction, which includes a record of immoral acts like stealing or lying. Whether immoral actions in general can foster positive moral development is a question that goes beyond the investigation of love.
worse than the friendship between moral saints, and maybe moral saints will never be friends with them; but it is, nevertheless, friendship.

Because morality is not a necessary feature of friendship, there is no reason to tie love with universal or moral standards of well-being. Sometimes we become friends with people who do not foster our well-being understood in accordance with such standards—Louise’s well-being is not objectively being promoted when she plunges to her death. Sometimes we change our friends for the better—when Louise becomes more self-confident—, but other times we make them worse—when my friend cements my interest in gambling. Sometimes, the qualities of a loved person are morally commendable traits, but sometimes they are not—think of the people, especially women, who love serial murderers and send them letters to jail.

Do we, as Helm says, fail in our concern towards the loved person if we just ‘go along’ with their damaging values? We probably do, if concern is understood as promoting their well-being according to universal or moral standards. Can we still say we are in love with them? I have shown we can, given that such concern may be what makes the best friendships, but the lack of it does not invalidate that people can be friends in spite of it.

The mutual shaping view avoids Helm’s objections and offers a more suitable account to explain non-harmonic love than the caring view. However, it is still open to another objection—one that Cocking and Kennett recognise and that I do not believe they fully respond to.

4.1.4. Does love justify itself?

Helm makes a third objection to Cocking and Kennett: there is nothing in the view which reveals that friends shape us in a way that other people do not. “[W]e might just as easily accept such direction and interpretation from a mentor or even a stranger”, Helm points out (2010a: 260). The only way to solve this problem is, according to Helm, appealing to the rational commitment love entails: “We ought to accept direction and interpretation from our friends as providing us with defeasible reasons precisely because they are our friends” (Helm, 2010a: 260).

Perhaps puzzlingly, Helm omits Cocking and Kennett’s position on the matter. Although briefly, they explicitly address that worry. Cocking and Kennett recognize
that self-concepts are influenced by people who are not our friends: from parents to bosses and therapists (1998: 523). A therapist might tell John that he always wants to be right, and John may take that interpretation on board in the same way he takes his friend’s interpretation on board (see §4.1.1). Cocking and Kennett argue that the direction and interpretation in the context of friendship is “different in kind” (1998: 523) from the therapy case:

In the therapeutic relationship my acceptance of direction would not mark my developing friendship with the therapist but, rather, my belief that such direction is good for me and will help my recovery. Direction and interpretation in therapeutic relationships is characterized by the end of restoring the patient’s health. Once this end has been achieved it is appropriate to end the relationship. No such terminating condition applies in friendship. The direction and interpretation in friendship is not undertaken with any particular end in sight, let alone the end of psychological health. (Cocking and Kennett, 1998: 523).

For Cocking and Kennett, direction and interpretation in friendship does not have an end, understood as a purpose. This is the same as saying that the loved person’s direction and interpretation of the lover is done for its own sake (i.e. for no other purpose or end). On the side of the lover, she accepts the loved person’s direction and interpretation insofar it ‘marks her developing friendship’ with the lover. Although this answer avoids Helm’s objection, it opens a new one: it seems that love justifies itself in the mutual shaping view.

That worry becomes even more apparent in cases of conflict. Lydia and Mema may negotiate their conflict by reaching an agreement, like going hiking on alternate weeks. When I first put forward this example in §1.3.3, I said that theirs is plausibly a case of being in love, since it is out of love that they negotiate their conflict. But then, the mutual shaping view gets to an infinite regress. It is because Mema and Lydia love each other that they reach an agreement which mutually shapes them, but in order to be in that situation, they should already love each other—they should already be mutually shaping each other. Even though I am not concerned with the justificatory reasons of love in this thesis, I cannot argue for a view which justifies itself. The only option, then, is to return to Helm’s claim that we ought to mutually shape each other because love provides with defeasible reasons. Nevertheless, this
move need not follow Helm’s route on rational commitment. In the next section, I show an alternative route which is not care-based, but meaning-based—based on the narrative nature of love.

**4.2. Love as a trajectory**

In §2.1.2, I said that Helm correctly identifies another feature of love: love is historical. I now develop that feature drawing from Karen Jones (2008), who is, to my knowledge, the only philosopher who has a theory specifically about being in love—a theory which opens a way out of the objection I set above for the mutual shaping view.

Jones argues that love is not a feeling. I briefly mentioned her view on the matter in §2.1.1: love correlates with many different feelings. The feelings which arise when being in love, Jones says, can even be mistaken for a stomach ulcer (Jones, 2008: 268–269). According to Jones, what differentiates love as distinct an emotional phenomenon is that, unlike ulcers, love partly depends on interpretation (2008: 270). That is, one’s interpretation of certain feelings, attitudes and actions as being in love are constitutive of being in love. Although this idea, she explains, has “gone out of fashion” (Jones, 2008: 270), Jones says that defining the role of description and interpretation in emotional phenomena is the route to explain “higher cognitive emotions” (2008: 271) like love, guilt, regret, or remorse.30

Now, before getting into more detail about Jones’s account, I need to pause for clarification. In light of the above, it may seem that Jones’s view is out of place here, since hers is a theory of emotion. I already explained in chapter 1 that this thesis is not concerned with the metaphysics of emotional phenomena. However, Jones’s view is still useful for my purposes:

I want to revisit the question of the role that descriptions of our emotional experience have in our affective lives, not just because I am interested in the emotions, but also because I think that many properties

---

30 Jones mainly engages with the dialectic between Taylor (1985) and Moran (1988, 2001) on the role of interpretation in emotion. The idea that description and interpretation are constitutive of emotional phenomena is also widely endorsed by feminist philosophers, especially regarding anger (see Scheman, 1980; Bell, 2009; Srinivasan, 2018). These, however, pertain to the philosophy of emotion, which I have departed from—although I will partly support the feminist claim that social construction plays a role in defining what being in love is in §5.2.2.2.
that are of interests to moral psychologists, including practical-identity properties such as “queer” or “political activist”, display the same openness to future contingency and to our interpretations of them that love does. (Jones, 2008: 271)

Jones calls ‘practical-identity properties’ what I have been calling characteristics of the self-concept. Being queer is not an emotion; neither is being a political activist. They are practical identities (a term which I stopped using given that it is not often discussed at the level of the self-concept). Eventually, my view is that being in love is a characteristic of the self-concept. So it is in this sense that Jones’s view is relevant for my discussion.

According to Jones, being in love it what she calls ‘a trajectory-dependent property’. In order to understand that term, first we need to know that a trajectory

is an ordered, temporally extended sequence of states or events, where the kind of ordering required as well as the typical temporal length of the sequence depend on the kind of trajectory in question. (Jones, 2008: 271)

Also, whether a trajectory-dependent property applies at a given moment depends on past and future events:

[A]scriptions of trajectory-dependent properties have… temporally extended truthmakers such that… whether it is correct to ascribe a trajectory-dependent property to A at t depends on what happens elsewhen, whether at t+n or t–n. (Jones, 2008: 272; emphasis in original)

To see this idea more simply, Jones gives the example of the process of building a house. Someone’s hammering a nail to a wall only counts as ‘building a house’ if it is embedded in a trajectory which follows a sequence (laying down the bricks, setting the cement, painting the walls, hammering a nail, etcetera). Before the hammering of the nail, there must have been other actions connected to each other of which the hammering is a culmination (lay down bricks, build the walls, place the beams and then hammering the nail). ‘Building the house’ is a trajectory; the house’s property of ‘being built’ is trajectory-dependent.
There is a type of trajectory which, according to Jones, is “interpretation-sensitive”. This is best explained with examples. Take the trajectories of ‘being on a journey’, being on a quest’ or ‘being on a pilgrimage’. These three examples, Jones adds, are interpretation-sensitive, because whether one classifies one’s travelling as a journey, a quest or a pilgrimage will depend partly on how one conceptualizes the activity. Conceptualizing one’s travelling as a quest entails that one applies certain meaning to one’s travelling: for example, it has to be a journey of certain difficulty and it has to be a journey with a certain mission in mind. Once one has conceptualized one journey as a quest and endorsed such interpretation, the interpretation itself is a reason to continue in the quest. Jones claims that conceptualizing a journey as a quest already gives you certain reasons to continue travelling: “agents that conceptualize and endorse their activity under that description are more likely to bring about that the resultant trajectory meets the conditions than those who do not” (Jones, 2008: 274).

So far this seems mostly as a discussion on the metaphysics of processes that might seem of little assistance to solve the problem of conflict for the mutual shaping view. But the setting up is necessary to understand Jones’s claims about love. Like quests, love is an interpretation-sensitive trajectory:

> Love… is an ordered temporally extended sequence of events and states, where conceptualizing one’s experience under the description “love” and believing that response to be warranted is likely to bring about that one’s experience in fact has the kind of shape needed to count as love. (Jones, 2008: 274)

In virtue of this, Jones identifies ‘being in love’ as a trajectory-dependent property:

> “[B]eing in love” is a trajectory-dependent property, and particular thoughts, feelings, and actions count as the thoughts, feelings, and actions of lovers in virtue of their assuming a place in a broader unfolding entity with the right structure. (Jones, 2008: 274)

---

31 Here the emphasis is on partly, since the classification of travelling as journey, quest or pilgrimage will depend on non-interpretative facts too. For example, a journey cannot qualify as a pilgrimage if it is not aimed at a religious or spiritual destination, or at least is grounded on a religious or spiritual motive.
Jones claim rests on the assumption that ‘being in love’ is a conceptualization which is, to an extent, agent-dependent. What for some people is expressive of being in love, she argues, is not expressive of being in love for other people (Jones, 2008: 275). That is a crucial feature of being in love which will be one of the foundations for my own view. However, in order not to deviate from the discussion, I am going to accept this assumption and continue in my aim of showing how Jones’s view opens a solution for the mutual shaping view (I argue in favour of being in love as agent dependent in chapter 5). I then leave unexplained –for the moment– what ‘the right structure’ of a love trajectory may be and call it ‘romantic trajectory’ to capture that it has to be a trajectory with a specific structure.

The relevant point here is that, in virtue of being in a trajectory, one is more likely to act in a way that the trajectory is continued. Here, we have a plausible route to solve the problem of the mutual shaping view. The fact that two people conceptualize their experience as being in love partly depends on the fact that their present experience is embedded in a past series of events and states, that they conceptualize as being in love with each other. It is also dependent on a projected series of events and states in the future. Jones rightly points out that it would be upsetting to hear one’s lover say “I have loved you” if it is not followed by something along the lines of “for many years, and love you still” (Jones, 2008: 287, n.20).

Hence, Lydia and Mema’s desire to negotiate their conflict arises from the very fact that they are in love, given that the conflict happens at a time when the trajectory-dependent property of being in love with each other obtains. Negotiating their conflict is the response that is likely to bring about that their experience has the shape of love. This is not only because they have been in a relationship (i.e. their past events and states) but because they project the continuation of the loving trajectory into the future. It is easier to see this point by looking at cases of conflict where the partners do not engage in negotiation to solve conflict. Imagine that Lydia said that she is going to the football every week, no matter what Mema says; or that Mema gives Lydia an ultimatum (“Manchester United or me!”). The fact that they do not negotiate the conflict will not, by itself, disrupt the trajectory. But it is definitely not the response that will most likely allow the trajectory to continue into the future: negotiating the conflict is. It is not then the case that love is a justification for itself. The fact that the lovers conceptualize and endorse a trajectory gives them a reason
to bring about the continuation of it. If love consists in mutual shaping, then the fact that love is a trajectory in the sense described by Jones gives the lover a defeasible reason—to use Helm’s term—to accept each other’s direction and interpretation.

However, this view is too obviously deterministic. It would be overly ambitious, and most certainly wrong, to claim that once one is in a trajectory (be it a quest, or being in love), one would be irrational to exit it (taking the first flight back home or breaking up). The trajectory view only becomes reasonable once we incorporate Jones classification of interpretation-sensitive trajectories as narratives.

A narrative, as seen in chapter 2, is a representation of a series of events where these events make sense in light of each other. It makes sense to see trajectories as narratives, given that whether trajectory-dependent properties obtain at \( t \) depends on what happens at \( t-n \) and \( t+n \). That is similar to say that events in a trajectory make sense in light of each other, given that Jones says that interpretation-sensitive trajectories are “sense-oriented” (2008: 275). This means that some events will be more meaningful than others in the context of a trajectory. The action of taking the first flight home and giving up one’s mission does not make sense in the context of a quest. It is not a matter of one’s will being constrained by the quest, of one being rationally committed to quest-related activities. There is, however, a reason to not taking the first flight home when one is in a quest: that is, it does not make sense in the context of a quest. Let us see how this view fares in the case of conflict in love, and whether it solves the objection that love justifies itself.

When Lydia is thinking about what to do regarding Mema and Manchester United, she is understanding the current event—the conflict—within a trajectory: the narrative of their relationship. In the midst of an argument, Lydia may imagine herself storming out to go see a match, leaving Mema at home upset. This imagined action will have certain evaluative and emotional import for Lydia—she may judge it cruel, and she may feel a pang of guilt at the prospect of hurting Mema. She may thus interpret this action as not being characteristic of the shape of a love trajectory, given that their romantic trajectory is characterized by mutual respect and kindness. Lydia imagines this future possibility as part of the love trajectory she is currently in and she judges it not conducive to its continuation (since it is not characteristic of the shape of love). The action of storming out is intelligible in general but does not make sense
as a loving action in the context of their romantic trajectory. Thus, Lydia quickly discards that possible action.

It can be the case, however, that Lydia does not stop to reflect on what is happening and does storm out. Lydia is not volitionally or rationally committed to loving actions. Still, she can reflect on what has happened afterwards and formulate identical mental states. She will go through the events and, in telling them to herself as a narrative, arrive at a judgement that she has done a cruel thing, a feeling of guilt and the interpretation that her actions were not characteristic of the shape of a love trajectory. Lydia will then apologize to Mema and try to make amends, that is, rectify her action, which is disruptive for the love trajectory. This reveals something that Jones overlooks: trajectories per se do not guarantee that the people embedded in it will more likely do what is conducive to their continuation, but that they are likely to prompt a correction response if they perform a trajectory-disruptive action.\textsuperscript{32}

The example shows that an action at $t$ in the context of a romantic relationship narrative acquires its meaning in function of past events at $t-n$ and future projections at $t+n$. It is because Lydia and Mema are co-protagonists the relationship narrative that the romantic trajectory is that storming out does not make sense as a loving action. The lovers’ actions are thus incorporated into the trajectory, further shaping their meaning as romantic trajectories. These actions are then both meaningful in light of the trajectory and ‘sense-making’ with regard to the trajectory. In the context of mutual shaping, the romantic trajectory justifies mutual drawing and interpretation of the lovers, given that it grounds their trajectory-related actions.

The trajectory view complements the drawing view, but still does not solve the objection I presented in §4.1.4. The fact that storming out does not make sense in the context of a romantic trajectory is, indeed, a reason to not storm out, and a reason for negotiating the conflict instead. However, this is not very different from Cockings and Kennet’s claim on why the therapist’s direction and interpretation is different from a friend’s direction and interpretation. Let us remember, they give two reasons: that in love mutual shaping is done for its own sake and that it is a feature

\textsuperscript{32} This claim is compatible with the fact that Lydia has other, non-trajectory-related reasons to apologize; mainly moral reasons (the obligation to right a wrong or the correction of the vice of cruelty). I do not intend to claim, and I do not think Jones does either, that every action embedded in a trajectory is justified by the trajectory only.
of a developing loving relationship. In Jones’s view, what is the end of a trajectory? It seems that a trajectory is equally fulfilled for its own sake. What is the reason for acting in ways that continue the trajectory? That trajectory-disruption actions do not make sense within the trajectory. Although the notion of narrative offers a reason for mutual shaping –that mutual shaping is done within a trajectory and certain actions make more sense than others– it is still the case that love justifies itself. In order to solve the objection and avoid having to go back to rational commitment, the reason to act in a way that continues the loving trajectory has to be external to the love itself. Otherwise, the account ends eventually justifying itself. The place to find these external reasons, I argue, is the self-concept –the narrative self-concept which I have described in chapter 3.

4.3. Mutual shaping, narrative and the self-concept

The main claim of the mutual shaping view, i.e. that the people we love shape our self-concept, is commonly accepted in narrative theory. This claim is discussed mostly in relation to selves or personal identity, given that it is mostly Strong Narrativists that pay attention to the relational shaping of personal identity (Taylor, 1989; Schechtman, 1996). It is, however, fully compatible with Minimal Narrativism, which as I showed in chapters 2 and 3 is mainly concerned with self-understanding without the need to make claims on selfhood or personal identity. Doug McConnell illustrates the role of mutual shaping within the framework of the self-concept:

All people are vulnerable to having their self-concepts shaped by others… [P]eople depend on others to develop and maintain skills of self-narration and they are vulnerable to having the content of their self-narratives co-authored by others. (McConnell, 2016a: 29)

Narrative theory, then, acknowledges that self-concepts are shaped by other people, in the sense that self-narratives are partly co-authored. When Cocking and Kennett say that friends interpret and draw one another, they are making a similar claim. Given that John’s friend Judy tells him that he always likes to be right, John ends taking himself less seriously: “John’s character and his self-conception are also, in part, drawn, or shaped, by his friend’s interpretations of him” (Cocking and Kennett, 1998: 505). However, Cocking and Kennett do not fully account for the feedback
loop between self-concept and action—and thus the scope of the influence of love in the shaping of one’s self-concept.

4.3.1. Meaningfulness and salience in mutual shaping

When Cocking and Kennett discuss interpretation and drawing, it is explicit interpretation that contributes to shaping the loved person’s self-concept. The lover observes the loved person’s actions, formulates a judgement and communicates it to the loved person, who subsequently sustains or modifies her actions (and self-concept). It also seems to be publicly displayed action that draws the lover’s action. It is because the loved person is interested in ballet and publicly shows to the lover her interest in ballet (by telling her, asking her to come to a ballet show, talking frequently about ballet, etc.) that this draws the loved person towards going to the ballet. However, lovers can also interpret each other without formulating it explicitly, and direct each other without publicly displayed action.

I now bring back the view on the self-concept I developed in chapter 3. The self-concept is composed of defining traits or characteristics which are the expression of a life worth living. Defining traits articulate the view a person has about herself, others and the world. I showed that there is a feedback loop between self-concept and action, grounded in meaningfulness and salience in the two sides of the loop.

On one hand, the self-concept is expressed in action, which forms patterns over time. These patterns contribute to the stability of the self-concept. The way these patterns make their way into the self-concept is by connecting events in one’s life. Given that the connection of events is made through selection and interpretation, and that these events make sense in light of the other, the self-concept is narrative. That is the role of meaningfulness. What I added to this traditional view on the narrative self-concept is that having a certain self-concept makes certain courses of actions more salient than others. As I explained, while episodic self-understanding denies the role of meaningfulness, it cannot deny the role of salience.

Mutual shaping, as formulated in §4.1, accounts for the role of meaningfulness. When John incorporates Judy’s interpretation of his behaviour, he acquires self-understanding. In other words, he acquires a new meaning for the events in his life; a meaning he was oblivious to. With her observation about John’s actions, Judy selects certain events in John’s life (all the times he is trying to prove he is right) and
feeds John with that information. Judy interprets these events as annoying, or damaging, or inappropriate—something that, in the mutual shaping view, John can only gather from her explicit judgements. Armed with Judy’s selection and interpretation, John looks at these events and puts them together as pertaining to the thread “I am always trying to be right”. Trying to avoid this behaviour, he begins changing his reactions and conversations, purposely trying to be more open to other people’s opinions. Eventually, these actions form patterns over time and feed to his self-concept: as a result of Judy’s interpretation, he ends taking himself less seriously, diminishing what used to be a trait of his self-concept. In essence, what Judy does for John is not different from what the disillusioned philosopher of my example in §3.3.2 does for herself (we may recall, she looks at what she does and what she feels about certain philosophy-related tasks and concludes that being a philosopher is no longer a defining trait of hers).

However, given its inward-focused view on the self-concept, the mutual shaping view does not accommodate for salience. I say it is inward-looking because, as formulated, mutual shaping only influences what John thinks about himself. This view on the self-concept overlooks that self-concepts articulate a person’s view on the world. Mutual shaping, I argue, changes how we understand the world in general by changing our patterns of salience.

This is implied in Judy and John’s example, but remains unarticulated by Cocking and Kennett. When John takes Judy’s interpretation on board, actions and attitudes of his which had until then remained unnoticed become salient. Cocking and Kennett do mention salience: “beyond making salient an existing trait of character, the close friend’s interpretation of the character trait or foible can have an impact on how that trait continues to be realized” (Cocking and Kennett, 1998: 505). However, that alone does not do justice to the relevance of salience in the process of mutual shaping, because all they mean is that by explicitly interpreting John’s behaviour (i.e. telling him “you always try to be right”), Judy is revealing to him a feature of his self-concept. What I am trying to say here has a wider scope: a loved person’s interpretation makes salient features about the world which feed into the lover’s self-concept. Once Judy has made the trait of ‘always wanting to be right’ salient, John will start finding some of his own actions salient. It will become salient to him that he is arguing a bit too much in a conversation with a neighbour, trying to be right. It
will become salient to him that he was speaking over his employee when she was voicing a disagreement, trying to be right. Also, Cocking and Kennett do not acknowledge that it will become salient to John how other people behave with each other in relation to Judy’s interpretation. John may notice that his mother also tries to be right all the time. He may observe, for the first time, that Judy always tries to be charitable to his views – she is not always trying to be right. These observations regarding the characteristic of always wanting to be right will plausibly make the way into his self-concept. But this is only possible because Judy’s interpretation has established a new pattern of salience for him.

Another reason mutual shaping needs to accommodate salience is that loved people do not change our patterns of salience just as a result of explicit discussion. It can be the case that Judy never actually tells John that he always wants to be right, but nonetheless, this becomes salient to him by the mere fact that he shares his time with Judy and observes that she behaves differently. Observing Judy being a good listener and being charitable, and realising that this is one of the things he likes about Judy, may have the exact same effect as Judy’s explicit interpretation. We often look at people’s actions, and this observation sometimes results in new patterns of salience which subsequently change our own actions and self-concepts. An account of mutual shaping that includes salience can accommodate this as part of friendship; an account that only focuses on meaningfulness cannot, because meaningfulness is only obtained through explicit discussion of one’s actions.

4.3.2. Meaningfulness and salience in trajectory dependence

The trajectory view also makes the mistake of focusing solely on meaningfulness and ignoring salience. As I noted in §4.2, what makes Jones’s view on being in love relevant beyond the philosophy of emotion is that she also thinks that practical identities are trajectory-dependent. Practical identities, as I understand them in this thesis, are defining traits of the self-concept which are expressions of one’s conception of a life worth living. Jones does not go on to elaborate on her practical identities examples of being queer or being a political activist, but I do so here. Because of the complications that come with defining queer identity and my own reticence to make claims on such topic as a non-queer person, I am going to use the political activist example. A person can do an isolated act of activism, i.e. joining Extinction Rebellion’s die-in on a road of London as a protest against climate change
inaction. But she can only call herself a political activist if she either has a past history of protests or she continues joining protests in the future. Imagine an adult millennial telling you she was a political activist in the early 2000s. “What did you do?”, you may ask. If she answers “I went to an anti-war demonstration once”, you would be right to point out to her that she was not a political activist at all. Certain actions only define a person as a political activist in function of what has happened elsewhen, to use Jones’s term. The practical identity or characteristic of being a political activist is only intelligible as such in virtue of being embedded in a trajectory, i.e. it is trajectory-dependent.

Saying that being in love is a ‘sense-making’ trajectory-dependent property accommodates meaningfulness but not salience. I have said above that Cocking and Kennet’s view is too inward-looking because it focuses solely on how mutual shaping influences a person’s beliefs about oneself. Now, I say that Jones’s view is too outward-looking, because by focusing solely on how the actions are conceptualized with respect of the trajectory, she overlooks how the trajectory—the relationship narrative— influences a person’s beliefs about oneself, i.e. her self-concept.

The trajectory of the anti-climate-change political activist will give a different meaning to different courses of action, and some actions like accepting a job for a fracking company do not make sense with respect to the trajectory. This mirrors Lydia’s storming out in the middle of an argument, which does not make sense with respect to that action being embedded in a love trajectory. However, alluding only to trajectory does not account for the patterns of salience that being a political activist or being in love trigger. The political activist is going to find certain things salient in virtue of being a political activist: news about the heatwave, signed contracts about fracking, the fact that she sees fewer and fewer bees every year. Someone who is not a political activist may find these facts less salient, or more probably not salient at all. Lydia is going to find certain things salient in virtue of being in love with Mema: hiking-related Facebook events, job offers closer to their residence, the need to find a sitter for her dog if they go on holiday together. If she was not in the specific trajectory that a loving relationship with Mema embeds her in, she would find these things less salient, or more probably not salient at all. Lydia finds certain things about the world salient in virtue of being in love with Mema.
Furthermore, the trajectory view overlooks that Lydia will find certain things about herself salient in virtue of being in love with Mema, on the basis of my earlier claims on how mutual shaping influences patterns of salience. As I said, being in love with Mema will feed into her self-concept both through Mema’s explicit interpretation and direction and her own interpretations about herself when observing Mema.

4.3.3. Justifying the shift to the self-concept

The mutual shaping and the trajectory views are on the right track to explaining what being in love is. But without incorporating the role of salience, they face the objection of love justifying itself. There is a way out of the objection by moving the discussion of trajectory-dependence and mutual shaping to the framework of the self-concept.

Once again, let us remember two features of the narrative self-concept I described in chapter 3. Defining traits of the self-concept are expressions of the life one finds worth living. Also, self-concepts retain consistency over time through a feedback loop with action, grounded in meaningfulness and salience in both ends. If being in love is a characteristic, then there is an independent reason to accept the loved person’s direction and interpretation. Given that self-concepts are grounded in what a person finds worth living, one wants to continue the trajectories that make her life worth living. Like, for example, being a political activist. Or being in love, which has as a necessary condition that one accepts the loved person’s direction and interpretation.

When one is in love with someone and endorses being in love with them, one wants to continue being in love, because being in love with that person is part of what one considers a life worth living. Imagine Lydia being told that tomorrow she will wake up and she will no longer be in love with Mema. Her relationship with Mema is a component of a life worth her living, so that would be a horrifying prospect. Lydia knows that not being in love with Mema will fundamentally change her life in a way she does not want it to change. Thus, she has an independent reason to continue in her loving trajectory—a trajectory where she accepts Mema’s direction and interpretation. This view does not rest, like Helm’s, on a rational commitment. Self-concepts change, as I showed in the case of the disillusioned philosopher. People fall out of love. A rational commitment to love—be it a pattern of emotions or a
trajectory– is unsustainable. But the view of being in love as a defining trait of the self-concept incorporates all the right conclusions of the mutual shaping and the trajectory views without love being a justification for itself – and without the need for a rational commitment that is incompatible with non-harmonious love and, ultimately, unsustainable. I develop this approach in the next chapter.

Conclusion

The mutual shaping view, according to which the lovers direct each other’s actions and shape each other’s self-concepts, overcomes the problems of the caring view while accounting for its right claims about love. Mutual shaping explains how some of the loved person’s interests become the lover’s interests while accommodating non-harmonious love. However, the mutual shaping view cannot explain why lovers accept each other’s shaping: it seems that, without a defeasible reason, love justifies itself. The trajectory view offers an independent defeasible reason for love. Being in love is a trajectory-dependent property, which obtains in virtue of it happening in a loving relationship. The reason to continue in the trajectory is given by its narrative structure: some actions make more sense as pertaining to the loving trajectory, while others do not. The lovers are not then rationally constrained by love, although they are more likely to take the actions which continue the trajectory. I argued that although the trajectory dependence of being in love does offer a reason to continue on the trajectory, it is not still a defeasible reason. Even after combining the mutual shaping view with the trajectory view, love still seems to justify itself.

I then showed the shortcomings of both the mutual shaping and the trajectory views. Mutual shaping is inward-looking insofar as it only explains changes in the lover’s understanding of herself, and trajectory dependence is too outward-looking insofar as it overlooks the changes that come in virtue of being on a trajectory. Both views are grounded in meaningfulness without accounting for salience. I have argued that salience is an essential component of mutual shaping, given that being in love changes one’s patterns of salience, not only about oneself, but about the world. Understanding love as a trait of the self-concept accommodates salience, and also gives a defeasible reason. Since defining traits of the self-concept are expressions of the life the lover finds worth living, the lover has an independent, defeasible reason to accept the loved person’s mutual shaping and to continue the loving trajectory. Not doing so would entail a loss of an element of the life she finds worth living –
being in love with the specific loved person. This view on being in love as a characteristic of the self-concept overcomes the problems of the mutual shaping, trajectory and caring views. Now that I have shown that this is the best approach to define what being in love is, I proceed to describe the view in detail.
Chapter 4

Being in love

Having justified the shift towards the self-concept, I can now develop in detail the claim I introduced in chapter 2:

Being in love is a defining trait of the lover’s narrative self-concept, which articulates the lover’s understanding of herself and the world.

In this chapter, I offer support for this claim by bringing together the conclusions I have extracted throughout the thesis about love and the self-concept, justifying some features of being in love that I have left unexplained, and responding to objections.

In §5.1, I explain what being in love is in the context of my account, which I call the ‘trait view’. I summarise the view I have introduced in chapter 4, according to which being in love is trajectory-dependent mutual shaping. Then, I show that understanding being in love as a defining trait reflects better the role of the loved person as a component of the lover’s conception of a life worth living.

In §5.2, I face an objection while at the same time explaining a feature of being in love: there are no agent-independent conditions for an instance of personal love to qualify as being in love. I argue that being in love is, to an extent, agent-dependent. Individuals conceptualise what being in love is for them through three levels of narrative: individual narratives, relationship narratives and socially shared narratives. I argue that this approach differs from a purely social constructivist view, which cannot accommodate off-the-mainstream plausible instances of being in love.

Finally, in §5.3 I briefly respond to a last objection on whether the trait view allows for self-proclaimed non-narratives—such as Galen Strawson—to be genuinely in love. My answer is that they can, although they can never obtain maximal understanding about what is, for them, to be in love.

5.1. The trait view

5.1.1. Trajectory-dependent mutual shaping

With the discussion of the previous chapter in hand a plausible definition of being in love is that being in love is a trajectory-dependent property which has mutual
shaping as a necessary condition. I said in §1.4 that my aim in this thesis was to investigate the change that love represents for the lover. Once my criticisms to the mutual shaping and the trajectory views (§4.3.1 and §4.3.2) are incorporated, that way of defining being in love partly explains this change in the lover. I will summarise it again with an example of a couple, Emma and Seb.

In virtue of being in love with Seb, Seb’s interests will direct Emma towards certain actions; for example, going hiking. Emma will go hiking with Seb in virtue of that being one of Seb’s interests, and that may result in hiking becoming one of Emma’s values.

Also, being in love with Seb will make certain things about himself salient to Emma, in virtue of Seb’s interpretation. Seb may tell Emma that she is very funny, which is something that she was oblivious to. Emma’s trait of being funny becomes salient to her: it will now draw her attention that she usually has witty comments to make, and that friends always count on her to animate a boring party. In that way, Emma acquires a new way of understanding herself. As I said in §4.3.1, Cocking and Kennet only discuss explicit interpretation, but changes may be brought too by implicit interpretation. Seb may become silent and distant when Emma raises her voice, or may tell Emma that she raises her voice too much without explicitly telling her that he thinks she is an angry person. That may make Emma realise that she is an angry person, and she will not only acquire self-understanding, but eventually change that trait of hers in virtue of Seb’s interpretation.

In §4.4, I said that the mutual shaping view is too inward-looking because it accounts for the role of meaningfulness but it does not account for the role of salience. In line with the above, Seb’s interpretation of Emma as being funny will make certain things salient which are not things about herself, but about the world. For example, she will start noticing that Seb is always laughing when he is around her; she will notice how other people are bores at parties; or start paying attention to stand-up comedians who have a similar humour to hers. Also, Seb’s interests will become salient to her, and she will act accordingly. She will plan a hiking holiday instead of a beach holiday, for example. Finally, being in love with Seb will make some features about the world salient to her, i.e. hiking-related themes.
Being in love then changes one’s self-concept. As seen in §4.1.2, these changes cause new changes which require mutual dynamic adaptation. That was Rorty’s claim, which is also endorsed by Gary Foster:

> We must come to know another person in order to be aware of her interests, but we should never let this knowledge become frozen so that we maintain a fixed concept of the other. We must constantly allow our conception of the other to be transformed by her freedom as it is expressed in dialogue, communication, through living together, and experiencing each other. (Foster, 2008: 246–247)

A disruption in dynamic adaptation can potentially bring the extinction of love. As Cocking and Kennett argue, a lover cannot accept the loved person’s interpretations if they do not track the changes in her. Rorty also set this responsiveness to change as a condition for love. One way of understanding this condition for the continuation of being in love is to say that if the loved person does not respond to the changes in the lover, she is not in love anymore. However, this claim needs to be qualified.

Imagine that Seb loses his interest in hiking, but nevertheless, Emma continues booking hiking holidays and buying hiking gear for his birthday. This alone cannot be a reason to say that Emma is not in love with Seb. Demanding responsiveness in this way would mirror the mistake of other views based wholly on harmonic love. Surely, in harmonic love the lover responds adequately to all the changes in the loved person; but I have argued that a plausible theory of love should account for non-harmonic cases too. On the other hand, if Emma is completely unresponsive to all the ways Seb changes over time, it would be correct to say that she is not really in love with Seb anymore, since her description of the person she is in love with differs too much from Seb’s self-concept. Neither Cocking and Kennett nor Rorty foresee this problem, and the answer needs to be, as in mutual shaping, a broad one: the lover needs to be responsive to at least some relevant changes in the loved person. By ‘relevant’ I mean they have to be changes in the loved person’s values. It would not be enough if the lover only adapts to physical changes, or the loved person’s accepting a job in a different company, for example.

Openness and responsiveness to mutual shaping is then a necessary condition for being in love, but as I explained in §4.1.4, there has to be a defeasible reason for
mutual shaping that is not the love itself. One reason to accept the loved person’s mutual shaping is that being in love is a trajectory-dependent property.

As I showed with the example of the political activist, going to one demonstration once does not warrant you calling yourself a political activist: there must be more events of political activism before and/or after that isolated event. Equally, one is not in love for a day: that day of being in love has to be embedded in a trajectory of having been or going to be in love with a person (at least in the form of a projection).  

Being in love makes certain actions more meaningful than others as the actions which belong to a romantic trajectory. Imagine Seb repeatedly telling Emma that he has moved on from hiking and now he prefers to go to whisky tastings. If Emma were to ignore these remarks and continue booking hiking holidays and buying hiking gear for Seb’s birthday, that would not make sense as an action that promotes the continuation of the romantic trajectory. It would not necessarily mean that Emma is not in love with Seb, but it would be detrimental to the romantic trajectory. In virtue of being in love with Seb, it is more likely that Emma will eventually correct this action and buy a whisky tasting experience for his next birthday: she will do that because she is in love with Seb. Here is where we reach a point of circularity and there is a threat of love justifying itself.

The continuation of the romantic trajectory is ultimately justified by self-concept consistency. People have a tendency to promote the actions which contribute to the consistency of the self-concept, and that is a reason for acting in a way that continues the romantic trajectory. This tendency to stability is not in contradiction with the condition of dynamism: it is widely accepted that self-concepts are malleable (i.e. dynamic) but mostly tend towards stability (i.e. consistency) in psychologically healthy individuals (see Markus and Kunda, 1986; Markus and Wurf, 1987;  

33 A question may arise here about the conditions for the property of being in love starting to obtain. Elaborating on this point would require a deviation from the current discussion and, in any case, I think this is more a question about falling in love than about being in love. In any case, one possible answer which would be in line with my narrative account is that formulating the thought ‘I am in love with this person’ in light of what one feels and does, gives the loving trajectory what Doug McConnell calls “self-narrative momentum” (McConnell, 2016b: 208). Another route to answer this question is Richard Moran’s idea of self-interpretation requiring to “making up your mind” about what you feel about certain persons (Moran, 1988: 141-142), but that would require a discussion at the level of the self, not the self-concept.
Since Emma is in love with Seb, she will likely act in a way that is consistent with that trajectory-dependent property, or correct actions which are inconsistent with it. Self-concept consistency is, then, a defeasible reason for continuing in the love trajectory.

At this point, the view I have portrayed here may be criticised for being deeply unromantic. Although the objections I presented to caring view, which is too focused on disinterestedness, may have seemed convincing, it may seem that I have gone too far. Trajectory-dependent mutual shaping does clearly overcome Helm’s criticism of individualism, since it fully accounts for the claim that people’s identities are formed relationally, but it may seem to fail in his challenge of egoism. This objection, however, would follow from a very narrow understanding of self-concept consistency that I will challenge next.

5.1.2. Self-interest and a life worth living

Helm’s criticism of egoism is directed both to robust concern and union views. According to Helm, in these views, the lover only cares about the loved person’s interests in virtue of these having become the lover’s own interests. This, Helm says, is an egoistic conception of love. Now, it may be objected that the view on being in love as trajectory-dependent mutual shaping is also egoistic: that Emma only cares about hiking because hiking becomes part of her self-concept and not because that is part of what makes Seb’s life worth living. It could be argued that, in my account, Emma does not love Seb for his own sake – i.e. for himself and not for any other end.

To respond to that objection, I will repeat something that I have said before in this thesis (§1.4). The fixation with egoism in romantic love has led many philosophers to concentrate their efforts on demonstrating that romantic love is not self-interested (see Introduction and chapter 1). But romantic love is partly self-interested.

34 I have deliberately avoided to engage with the question of whether Helm is right in his egoism critique of Frankfurt. It is striking to find Frankfurt’s account being labelled as an egotistic view, given Frankfurt’s claim that in love “selflessness and self-interestedness coincides” (Frankfurt, 2004: 62). I remain agnostic on that front: my criticism to Frankfurt was not that rests on an egoistic conception of intimate concern, but that it only accommodates highly idealised examples of love.
Monique Wonderly (2017) differentiates between two types of self-interest in romantic love. The lover, Wonderly argues, may satisfy certain desires through the loved person—for example, culinary, sexual, or material desires (2017: 238). If the loved person’s ability to satisfy these desires are the only reasons for the lover’s being in love, then the loved person is fungible (Wonderly, 2017: 239). That is, she is interchangeable for anyone who can satisfy these desires, given that it is not who she is that is important, but her ability to satisfy these desires. This is indeed a questionable type of self-interest, but it is not the self-interest which is constitutive of personal love. Wonderly gives an example of a husband whose wife is ill, and gives the following statement on his hope for her recovery:

I do hope she recovers because she is a wonderful woman in her own right, but also, in part because without her, my life would be less meaningful—less fulfilled. Without her, I would not be able to get along in the world as well. And for a time at least, I would not be okay, but I would feel as though I were no longer all of a piece. (Wonderly, 2017: 239)

This is a second type of self-interest which is deeper than just interest for what a person ‘gets out of’ a relationship (like nice dinners or expensive cars). The husband in the example attributes to his wife a “particularly meaningful role in how he views himself and how he lives his life” (Wonderly, 2017: 240). This is a self-interested response to his wife’s value, but nonetheless a loving response (Wonderly, 2017: 240). This self-interest is not only permissible, but constitutive of love (Wonderly, 2017: 237).

To illuminate that claim, consider the following exchange between husband and wife: “I hope you recover because you are a wonderful woman in your own right”; “But how would you feel if I die?”, the wife could ask; “Oh, I would be sad because of how valuable you are as a person, but ultimately it would make no difference in my life”. As much as the husband values his wife’s identity as a whole (i.e. for her sake, ‘she is a wonderful woman in her own right’), his wife would probably think that her husband does not love her very much. What she would like to hear is plausibly something along the lines of “Not having you would leave a hole in my life that would make it hard for me to go on”. This idea corresponds with everyday
experiences of what losing a loved one entails, which has been captured in countless songs, stories and real-life reports.\textsuperscript{35}

My claim here is that this happens because, as Wonderly argues, the loved person contributes to one’s leading a meaningful life. In other words, the loved person is a component to what I have been discussing in the context of this thesis as ‘the conception of a life worth living’.\textsuperscript{36} If Emma is in love with Seb, being in love with him is part of her conception of a life worth living. Recall that a person’s conception of a life worth living is expressed in that person’s values. This is precisely what my view reflects: in virtue of being in love with Seb, Seb is part of Emma’s conception of a life worth living. It is then a value of Emma’s. I will unpack this idea further.

I bring back the counterfactual I offered in §4.3.3. Imagine Emma was told that she would wake up tomorrow and not be in love with Seb anymore. Emma would think that, were that to happen, she would lose a component of the life she conceives as worth her living. I said above that the egoism objection rests on a very narrow interpretation of consistency. It is not a matter of one choosing actions which are consistent with one’s self-concept, merely letting oneself go through the motions one’s characteristics have put one into. It is, instead, a person’s aim to maintain a life which is authentic: a life in accordance to that person’s conception of a life worth living, which, as seen in chapter 1, constitutes her identity. The way the people we love direct us towards that aim cannot, and should not, be discussed in equal terms of those same people satisfying our culinary, sexual or material desires, to use Wonderly’s examples.

If this view brings now the suspicion of been highly idealised, there are two responses to that line of objection. Firstly, think of the examples of values that I

\textsuperscript{35} Two good examples are singer-songwriter Harry Nilsson (“I can’t live, if living is without you”); or Emily Brönte’s (1847 [1992]) Heathcliff (“Be with me always—take any form—drive me mad! Only do not leave me in this abyss, where I cannot find you! Oh, God! It is unutterable!”).

\textsuperscript{36} Sam Shpall’s (2018) ‘tripartite theory of love’ is also based on the contribution of love to a meaningful life. I do not engage with Shpall’s theory here for two reasons. Firstly, it is a theory of love in general, including not only love for persons but also for objects and ideas. Secondly, it is a definition of what love is, while I am defining being in love as a component of romantic love—not what romantic love is. All in all, I am highly sympathetic to his view of love as “devotion that renders vulnerable and expresses liking” (Shpall, 2018: 91; emphasis in original). Shpall’s view will certainly feature in any future incorporation of the work I do here to a theory of what romantic love is. See also Susan Wolf (2015) for the role of love in a meaningful life –Wolf, however, defines love as a type of caring for the loved person for their own sake (2015: 189), an approach I have rejected in chapter 1.
have offered throughout this thesis. Imagine Manchester United was to disappear due to some economic or personal disaster. Incidentally, this seemed like a real prospect in 1968, when a plane crash killed half of the team and left many of the survivors thinking that they would not be able to play again. This is how a lifelong Manchester United supporter describes how he felt when the news broke:

> When the headlines appeared on placards outside the newspaper shop, I knew it was true. It was one of the worst moments of my life. I felt the same kind of sadness as when my dear old mother died. It took me ages to get over it. (Hughes, 2008, n.p.).

For the people that have Manchester United as one of their values, losing Manchester United would entail losing a component of what makes their life worth living. Losing one value does not mean that one’s life stops being worth living as a whole, but it is a loss in self-interest which is different in kind than other losses in self-interest (for example, a Manchester United fan may be devastated when the team loses a title, but that cannot compare to losing the team itself). This applies as well to values which do not conform to universal or moral standards: Doherty may, if he eventually overcomes his drug addiction, feel like he has lost a part of his life worth living by giving up drugs. The mafioso may feel like he has lost a part of his life worth living if the police bring down his crime syndicate. This is not a highly idealised view on what being in love is: it reflects the fact that that losing a loved person can have the same effect as losing other values can affect a person’s conception of a life worth living.

Secondly, so far I have only discussed cases where the lover endorses being in love with the loved person, so it may seem that the view only accommodates harmonic love. However, that is not the case. The view I am portraying here is compatible with the fact that sometimes we do not want to be in love with the people we are in love with, i.e. we do not endorse our being in love. This may happen when we are in love with people who are damaging to us, who do not love us back or who represent a significant conflict with other values that we hold. As seen in my discussion of problematic identities (§1.2.3), one’s values might be in conflict with the conception of a life worth living one believes one should have. However, they are still in accordance with the conception of a life worth living one has. Doherty’s conception of a life worth living as a life of sex, drugs and rock and roll may be perverted or, in
Helm’s terms, ‘somehow rotten’, but is nevertheless his conception of a life worth living. Equally, when one does not want to be in love with the person one is in love with, one has a desire for a different conception of a life worth living—but being in love with that person is nonetheless constitutive of it as things stand. My view then accommodates non-harmonic love, since, once again, I am not establishing what is the best kind of love, but offering a definition of being in love.

This inclusion of non-harmonic love reveals another reason why the view is not egoistic, besides being based on a specific kind of self-interest. Being in love with someone is part of one’s conception of a life worth living, even if that conception is objectively damaging to oneself. One can be in love with someone who is self-destructive and only brings destruction to one’s life—probably the case with some of Doherty’s partners. One can be in love with someone with whom every little gesture may unleash a political disagreement, making one’s life significantly more uncomfortable—like in the case of Mangan and her Tory husband. But one is, nevertheless, in love with that person, and no one else. That person, regardless of her moral worth or her conflicting values, regardless of whether she makes one’s life better or worse according to universal or moral standards of well-being, is part of what makes one’s life worth living.

By accepting that one is in love with someone and promoting that love, one does not know the direction one’s life and one’s self-concept will go. Recall Rorty’s dynamism: loving may determine where one ends up living, what one ends up liking and even the abilities one ends up having (Rorty’s example was Ella’s voice being potentially affected by the direction her life with Abe takes her). Furthermore, I think that there is a case to make for the fact that the beginning of loving relationships is equally exciting and frightening precisely because of this open-ended nature, the uncertainty of what is to come: one does not know how one will change by sharing one’s life with another. Nehamas says that friendship is “a commitment to the future, a sense that there is more to know here, and a promise that what I still don’t know will be worth learning” (2016: 133). But there is no guarantee that it will be, in fact, worth learning. It is because one loves a specific person that one is willing to accept that risk. In other words, this risk is accepted because who the other person is. I think this approach does justice to the importance of the loved person’s identity in a way that
the caring view fails to capture. It is not a matter of the lover valuing everything that the lover values, it is a matter of the lover *risking* what she values for the loved person.

On that basis, the objection of egoism to trajectory-dependent mutual shaping does not hold. However, the discussion in this section rests on the idea that being in love is a part of what makes a person’s a life worth living. In that sense, trajectory-dependent mutual shaping falls short of explaining how being in love fulfils that aim: it merely explains the mechanism that changes the lover, and the reason why the lover accepts these changes. The fact that the loved person is a feature of the lover’s conception of a life worth living is not explicitly accommodated by trajectory-dependent mutual shaping. That fact is accommodated if being in love is defined as a trait of the self-concept—which is my own view.

### 5.1.3. Being in love and the self-concept

In §4.2, I showed that for Jones being in love is an interpretation-sensitive property, in the same way other “practical-identity properties” such as being queer or being a political activist are (Jones, 2008: 271). What Jones does not say explicitly is whether she considers being in love itself a practical-identity property, to use her term. Here, I propose that being in love is indeed the type of ‘property’ that being queer or being a political activist is: a practical identity, which, in the context of this thesis, I have referred to as a defining trait of the self-concept (see §1.2.2 and §2.1.3 for my reasons to depart from Korsgaard’s terminology). I call this approach the ‘trait view’, which has the following basic claim:

Being in love is a defining trait of the lover’s narrative self-concept, which articulates the lover’s understanding of herself and the world.

The self-concept, I have argued, articulates a person’s understanding of the world through that person’s defining traits, which are expressions of that person’s conception of a life worth living. Throughout the thesis, I have mentioned several examples: being a philosopher, being a political activist, being an Instagrammer; also valuing drugs or gambling. I have also shown that a person’s defining traits articulate her understanding of herself, others and the world, by making certain things salient.

I bring back here Mary Jean Walker’s (2012) example of having the defining trait of being very organised that I offered in §3.3.2. Firstly,
someone who views him- or herself as very organized may tend to remember instances that show him or her this way, more easily and more often than instances that show the person to be disorganized. (Walker, 2012: 65)

Being in love has the same effect. To understand this in a simple way, think of a series of events that may happen to Rosa, who is in love with María. In the past month, Rosa and María had a romantic dinner when he felt that she really is in love with María. It was a lovely dinner, despite María complaining about the price of the dinner. Rosa and María usually go to lovely dinners, but they also tend to have arguments about money. However, Rosa tends to forget the arguments about money and remember the lovely dinners, because she is in love with María. Being in love makes a person select the events that are expressive of being in love more than the events that make her doubt that she is.

Going back to the trait of being organised,

how we remember events is affected by our view of ourselves and by the way we understand those events to relate to other events. For example, an event showing the person to be organized may be connected thematically to other similar events within their narrative—even if the same events could plausibly have a range of other meanings. (Walker, 2012: 65)

A lovely dinner, taken in isolation, could have different meanings: it can be a celebration, or a release of stress, or a treat. It can be then connected to other different events and thus, other traits. If one is writing a PhD thesis, the dinner can be a celebration of submission. If one is a workaholic, the lovely dinner may be connected to this trait as an exception of what one’s life usually is. If one is very frugal, the dinner takes the meaning of ‘treat’ in virtue of its connection with other events that express one’s frugality. But for Rosa, who is in love with María, the lovely dinner is connected to her being in love with María; the moments they share together as a couple. In these ways, being in love with María articulates her understanding of herself.
I have also shown how having different defining traits leads people to select and interpret different events in their understanding of the world. We saw an extreme example of this with Ziggy and Billy (§3.2.2). In less outlandish examples (§3.3.2), the philosopher found a metaontological puzzle the most salient event in her day, while the speechwriter found the sudden coherence of the Prime Minister the most salient event. The same goes for Rosa: finding her being in love salient not only influences her understanding herself (as being in love) but also her understanding of the world, through trajectory-dependent mutual shaping, as I have explained in detail in the previous chapter and I have summarised above (§5.1.1).

However, being in love is not merely a mechanism that makes one remember certain events and forget others, or give certain events a specific meaning, or make certain things salient. It gives her certain reasons to act, as I showed in §3.3.1:

[The self-concept] encapsulates our reasons for action, since these reasons are related to our understanding of our selves and our pasts. Thus, it frames and guides our decisions about how to act. (Walker, 2012: 65).

The defining trait of being in love gives Rosa certain reasons for action, and thus influences her choices. But then again, by itself the providing with reasons for action does not capture the relevance I have given above to the loved person as a feature of a life worth living (§5.1.2). I bring back now Daniel Hutto’s claim on the role of narrative from §3.1.3:

By providing a window on our actual and possible doings the narratives we fashion about our lives can help us to decide, for example, if our taking this or that action is something we want to figure in the story or stories that we would want told about who we are. (Hutto, 2012: 26)

This is the key reason why understanding love as a defining trait does better justice to the importance of the loved person for the lover than trajectory-dependence mutual shaping. For Jones, interpretation-sensitive trajectories make certain events more meaningful as pertaining to a romantic trajectory than others. That is something I have fully incorporated into my view. But I take it further now by recalling that interpretation-sensitive trajectories are narratives (§4.2). This narrative,
the romantic trajectory, is *one of the stories that the lover person wants to be told about who she is*. This is why the loved person is part of the lover’s conception of a life worth living. For Rosa, María is a feature of that story she wants to be told about herself. In virtue of mutual shaping, she is a co-author of it, and hence the loved person is not only part, but *co-creates* what is a life worth living for the lover.

With this account, the trait view incorporates the trajectory-dependent mutual shaping view while accounting for the important role of the loved person. The role I have given here to the loved person as a feature and co-creator of the loved person’s conception of a life worth living follows from trajectory-dependent mutual shaping, but is only properly articulated with the understanding of being in love as a defining trait. The trait view also incorporates the right intuitions of the caring view. Frankfurt says that caring about something means that one suffers when that thing’s well-being is diminished (§1.1.1). Helm says that caring about something is for that thing to have import in you, which entails a rational commitment to a pattern of emotions (for example, sadness when that thing’s well-being is diminished) (§1.2.1). Having the defining trait of being in love explains that the lover’s well-being is affected by the loved person’s, as well as the loved person’s sometimes feeling the same emotions as the loved person. But, unlike the caring view, the trait view does not define being in love through these features, given that such approach denies that non-harmonic love can be genuine instances of being in love. In the trait view, this is *an effect* of being in love, not what being in love *is*. It may be argued that my view is a sort of union view, which I briefly mentioned in §1.1.1 as the view that defines romantic love as a merging of the lovers’ identities. It is not a union view: it is not *the loved person* that becomes part of one’s self-concept, but *being in love* with that specific person. Helm, I recall here, argues that the lovers end up sharing a single evaluative perspective (Helm, 2010a: 260). That is not my claim. I do not think the lovers share a *single* evaluative perspective, but given that they both partly co-create their respective conceptions of a life worth living, and the effects that mutual shaping has in the lovers’ self-concepts, there is a *possible overlap* in the way the lovers understand the world. Again, this is not a requirement of being in love, or what being in love *is*: it is an effect of being in love.

I will finish this section with a brief discussion of another advantage of the trait view: it opens a credible route to explain what falling out of love *is*. This discussion acts
both as further support of my view and as a bridge to the last argument of the thesis, on the necessary conditions for being in love in the context of the trait view.

5.1.4. Falling out of love

Many, if not most, philosophers of love fail to consider how their views on love fare when people stop being of love –i.e. when they fall out of love. One notable exception is Niko Kolodny, who gives a detailed argument on when it is rationally appropriate to stop loving someone (2003: 164-167). Mine is not an argument on reasons for love, so I do not engage with Kolodny here. In the context of my view, to fall out of love would be to lose a defining trait: the trait of being in love. This loss results from a disruption in the feedback loop, when one’s action does not correspond with the trait of being in love or when being in love stops articulating the lover's understanding of the world.

The case of the disillusioned philosopher is a paradigmatic instance of losing a defining trait: the philosopher finds herself bored with metaontology, dreading academic talks, more interested in reading sociology, for example. Equally, if Emma starts falling out of love with Seb, she may find herself bored with him, dreading their shared hikes, more interested in her new colleague. If Rosa starts falling out of love, she will start remembering with more frequency the arguments with María than the lovely dinners. The lovely dinners will be connected with other events: they now become an exception in her conception of their life together, or an example of what she would like to do with someone other than María.

The trait view does not only have a plausible way to explain falling out of love. It also accommodates the fact that the changes brought to the lovers’ self-concepts through mutual shaping may survive the extinction of love. Emma might not be in love with Seb anymore, but she may still value hiking. Falling out of love with Seb will not make her go back to being an angry person. This is not so clearly available for the caring view: if the lover is rationally committed to a pattern of emotions towards the loved person without them becoming one’s personal values.

Also, the trait view does justice to the importance of the loved person as the one who is doing the mutual shaping. The loved person is not just a ‘shaper’ but part of what makes the lover’s life worth living; a feature of the story the lover wants told
about her life. Losing a way of understanding oneself and the world can be highly disruptive, like falling out of love someone is especially if one still cares deeply for the other person and wishes one was not falling out of love. This effect has been investigated by psychologists, who have documented that romantic break-ups are often accompanied by a decrease in ‘self-concept clarity’. By this is meant a decrease in “the extent to which the contents of an individual’s self-concept (e.g., perceived personal attributes) are clearly and confidently defined, internally consistent, and temporally stable” (Campbell et al., 1996: 141). It has been found that exiting a relationship has a special negative effect on self-concept clarity: “ex-partners must renegotiate their sense of self without the facets defined by the relationship, leaving their self-concepts less clearly defined (at least temporarily)” (Slotter et al., 2010: 148-149). In moral philosophy, Ami Harbin includes break-ups in the kind of experiences she calls ‘disorientations’: “temporally extended major life experiences that make it difficult for individuals to know how to go on” (2016: 2).

Moreover, the trait view accommodates the fact that one may still be very fond, continue caring for or even being a close friend with the person she used to be in love with. Kolodny takes note of this fact (2003: 167), which is discussed at length in the context of break-ups –not falling out of love– in Lopez-Cantero (2018). The argument there is focused on the reasons of love, but the conclusion I reach is relevant here. I claim that when one still cares about the person one used to be in love with, and still has a relationship of shared activities with her, one has entered a new relationship with that person (Lopez-Cantero, 2018: 700). The new relationship lacks “romantic expectations”: expectations that are “expressive of love within a romantic relationship” (Lopez-Cantero, 2018: 701). Following Sara Protasi (2014: 216), I claim that “[e]ach romantic partner expects different actions from their beloved as the expression of their love” (Lopez-Cantero, 2018: 700).

Those claims, which are made as a response to several rationality-oriented views on love, are not precise enough for the purposes of this thesis. I will offer next, in the

---

37 This is the starting point for Pilar Lopez-Cantero and Alfred Archer (2019), where we develop the nature and value of falling out of love as a disorientating process in more detail. I am grateful to Alfred Archer for agreeing to me citing our work prior to publication.

38 In the paper, I argue that break-ups entail a loss of romantic reciprocity, i.e. reciprocity in a relationship conceptualised as romantic. What I offer next is an explanation of how individuals conceptualise what being in love is which partly invalidate that claim on what is lost in a break-up, which would need to be reformulated in light of the new argument I offer here.
context of the trait view, a more detailed explanation on the difference between being in love and being a friend. This also responds to the need to fully explain what is to be in love in the context of the trait view. So far, the trait view does not seem to differentiate between romantic love and friendship. This stems from the fact that, besides mutual shaping, I have not argued for necessary conditions for an instance of personal love to count as being in love.

5.2. What makes this romantic?

Romantic love and friendship have traditionally been understood as different “modes” of love (Kolodny, 2003: 139); as different in kind. This separation between friendship and romantic love tracks back to the Ancient Greeks, who distinguished between four types of love: eros, philia, agape and storge. There are different interpretations of each of these kinds of love, but roughly, *eros* is romantic love; *philia* is friendship; *agape* is love for humanity and *storge* is familial love—a well-known description of these ‘four loves’ can be found in C.S. Lewis (1960 [2012]). Many people seem to make this distinction in everyday language. ‘I only love you as a friend’, someone may tell you; and you will understand that this person means that she does not love in a way that is romantic. However, the trait view, as I have described it so far, does not seem to pertain exclusively to being in love, which I consider a component of romantic love. Indeed, many of the views I discuss in chapter 4 are views on friendship, and I said that their claims should be understood as applying to romantic love too. It may seem, that my view is defective in two ways: i) it does not offer necessary conditions for being in love and ii) it does not distinguish between being in love and being a friend.

In what follows, I answer to those two challenges while at the same time offering the last piece of the trait view. I have not yet still explained what is a romantic trajectory. My position is that, apart from mutual shaping, there are no agent-independent necessary conditions for an instance of personal love to count as being in love (i.e. to be in a romantic trajectory). Thus, there are no universal differences between being in love and being a friend.

---

39 I am aware that the division between *eros* and *philia* is not equivalent to the contemporary everyday language division between romantic love and friendship. I do not intend to endorse here that romantic love is analogous to *eros* and friendship to *philia*, either. But the Ancient Greeks’ division of these as two different kinds and the usual connection between *eros* and passion and *philia* and fraternity offer enough grounds to make the comparison.
5.2.1. The role of sexual desire

There is an intuitive answer to the challenges above that which I am going to quickly reject: the difference between romantic love and friendship lies in sexual desire, which seems to be widely accepted as a necessary condition or at least a characteristic activity of romantic love (for example, Kolodny 2003: 149). Natasha McKeever (2017) summarises this generalised position as follows:

For some people, sex is the distinguishing feature of a romantic relationship; the difference between loving somebody as a friend and ‘being in love’ with them is often taken to be the presence of sexual desire. Furthermore, all other things being equal, regular sex is seen as a sign of a good relationship and it is not unusual for people to worry about their relationships, or even to end them, if they are no longer having sex or having sex only very irregularly (McKeever, 2017: 203-204).

As McKeever says, this is the case for some people, but extracting a necessary condition from this preference would be a mistake. For a start, sexual desire is different in kind from romantic love. Raja Halwani (2017) gives several reasons for this distinction; I highlight two of them here. Halwani points out that sexual desire need not involve love: one may be sexually attracted to someone one does not love (as it happens in casual sexual encounters) (Halwani, 2017: n.p.). Also, people who love each other may lose sexual desire for each other and nevertheless continue loving each other—anthing that happens in some long-term loving relationships (Halwani, 2017: n.p.). So romantic love and sexual desire should not be understood as the same phenomenon, since they are different in kind.40

Secondly, sexual activity is not exclusive of romantic love. Besides from casual sexual encounters, one can be sexually active with a friend in a relationship popularly known as ‘friends with benefits’. In a psychological study, Melissa Bisson and Timothy

---

40 Halwani also says that love feels different from sex. In loving relationships, love is emotionally expressed through “tender kisses and caresses, whispers of “I love you””; but sometimes sex “takes over” the love (“the act becomes simply animalistic and sexual through and through”, Halwani 2017: n.p.). This is a normative claim on the features of sexual and loving activity that I do not endorse. I think that it is plausible that “animalistic sex” is expressive of love for some people, and even more plausible that some people who are in love do not express it through “tender kisses and caresses”. Hence, I do not include it as a plausible reason to distinguish between love and sexual desire.
Levine (2009) conclude that while being friends with benefits raises the worry of sexual activity potentially damaging the relationship, it also comes with the advantage of having a sexual partner who is known and trusted (Bisson and Levine, 2009: 72). Friendship potentially adds value to sexual activity as opposed to causal sexual encounters, so then friendship with benefits should not be treated \textit{prima facie} as a corrupted friendship. I do not intend to give a full argument in favour of friendship with benefits –only to show that it exists.

Thirdly, some people who are in love with each other cannot have sex with each other. There are relationships where one or all the partners have an impairment that does not allow them to have sex—a physical disability, PTSD or even old age. It could be argued that, if these people were in ideal conditions, they would have sex with each other as the best expression of their being in love. But again, it would be unfair to say that in the absence of sex, these people’s experience does not count as being in love.

Fourthly, some people identify as asexuals. Broadly, asexuals do not experience sexual attraction, although this is expressed in different ways:

Some [asexuals] find the idea of sex generally repulsive; others find the idea of themselves engaging in sex repulsive; others are neutral about sex; still others will engage in sex in particular contexts and for particular reasons e.g. to benefit a partner; to feel close to someone; to relax; to benefit one’s mental health, and so on. (McKeever and Brunning, 2019: n.p.).

Despite this attitude to sex, asexuals can be in relationships which they categorise as romantic (McKeever and Brunning, 2019, n.p.). That is, asexuals can be in love. The notion of asexuality may be resisted, and one foreseeable objection is that it is not a valid sexual identity (like heterosexuality, homosexuality or bisexuality, for example), but a sexual disorder. It is not my aim here to argue for the plausibility of asexual identity, but to show that, if asexuality is accepted as a sexual identity, that is another

---

41 I am grateful to Natasha McKeever and Luke Brunning for making a copy of their paper available to me prior to publication.
reason to separate sexual desire from romantic love – if not, there are still three other reasons above to do so.

It could still be argued that even if sex is not a condition for romantic love, and even if sex can be a feature of friendship, romantic love without sex and friendship with sex are not the best kind of romantic love and friendship. Once again: my aim is not to establish what is the best expression of being in love, but what being in love is.

In light of the above, if being in love and being a friend are to be differentiated as two traits of the self-concept which are features of two modes of love which are different in kind, it cannot be on the basis of sexual desire and/or activity.

5.2.2. A three-tiered conceptualisation

The position I am going to defend here is that what being in love is can vary between individuals, so that there are no agent-independent necessary conditions for being in love. I start by showing how individuals acquire their conceptions of what being in love is and thus the necessary conditions for an instance of personal love to count as being in love.

In §4.2 I acknowledged, but did not develop, Jones’s assumption that being in love is, to an extent, agent-dependent, given that what for some people is expressive of being in love, she argues, is not expressive of being in love for other people (Jones, 2008: 275). Specifically, Jones claims that

[j]everyone agrees that an episode of love… comprises many diverse elements: There are thoughts, patterns of attention, desires, various feelings, and actions. Somehow, from among the relatively inchoate mass of elements that form a person’s inner life, and from among the wealth of things that she does, we choose some items and not others to group together as constituting or expressing her love for the beloved. For any item we pick as being part of the person’s love, or expressive of it, it is possible for someone to have an exactly similar item in their mental economy and yet not count as being in love. (Jones, 2008: 275)

Jones then turns to trajectory dependence to explain what makes an episode count as being in love is, i.e. the fact that it is embedded in a romantic trajectory. Jones
highlights that a person’s interpretation of a trajectory as romantic will partly rest on how it is conceptualised in her social environment (2008: 281), but considers explaining such conceptualisation outside of the scope of her argument (Jones, 2008: 275).

Helm explicitly states that his account takes friendship as a model (2010a: 4). Later, he briefly states that romantic love is “a form of friendship in which the friends form a plural person in which the joint conception of their relationship more or less centrally involves romance” (Helm, 2010a: 288). Helm conceives his account of intimate identification as pertaining both to friendship and romantic love, which, he says, are not different in kind:

[Although there are undeniable differences between romantic relationships and relationships of friendship…, we ought not to trace differences to a distinction in kind between the evaluative attitudes grounding each. In other words, there is a single more basic kind of personal love underlying each; the differences between these various kinds of relationships can, I believe, be understood in terms of the particular ways the parties involved conceive of and negotiate the details of the relationships that emerge from reciprocated personal love. Consequently, we should not try too hard to distinguish philia from eros in providing an account of romantic love. (Helm, 2010a: 4)

The details of what “romance” may mean, i.e. what are the concrete differences between romantic relationships and relationships of friendship, are left unexplained by Helm (who acknowledges the fact in 2010a: 4). However, as seen in the quote above, Helm also seems to imply that these differences may vary between individuals.

I agree with both Jones and Helm, but I do think that in order to provide with a convincing account of being in love, I cannot leave the details unexplained. That would be a significant gap in the trait view. Given that my account rests on the lovers acting in a way that promotes the continuation of a romantic trajectory, I should explain which actions belong to a romantic trajectory. My claim is that a person’s conceptualisation of being in love is shaped by three narrative levels: relationship narratives, socially shared narratives and individual narratives.
5.2.2.1. Relationship narratives

Imagine you are in a double date at a restaurant and the couple sitting with you spends all the dinner caressing each other’s hair, calling each other cheesy monikers and holding hands on top of the table, leaving them only one free hand to eat. After saying goodbye and getting in a cab, you turn to your partner, roll your eyes and say, with a mix of relief and exasperation: “Thank God we are not one of those couples”. Superficially, the anecdote reveals that different people have different ways of expressing their being in love. I will argue that different people actually have, as well, different ways of conceiving their being in love.

People who are in a personal relationship like romantic partnerships or friendships do things together. This is a basic feature of the trait view and also of views which focus on the shared agency aspect of love (Helm 2010a; Nozick 1989), or on the relationship as a source of reasons for love (Kolodny, 2003). What usually receives less attention is that people in romantic partnerships and friendships, talk and think about what they do together, and therefore formulate a conception of the type of things they do together (such as caressing each other, or not). These discussions extend back to the shared past and forward into the projected future. Shaun Gallagher and Deborah Tollefsen call these ‘we-narratives’: “narratives about what we are doing, have done, and will do, or what we ought to do and want to do” (2019: 214). According to Gallagher and Tollefsen, the communication of a we-narrative is a source of stability:

First, it specifies what we are mutually assuring ourselves about what we are committing to. That is, the narrative specifies the content of commitments and assurances. Second, explicitly or implicitly, with respect to means and norms, narrative may indicate how we are going to proceed. Narrative can express both means and goals of future actions, and, although not as formally as a legal contract, agreement with the narrative can signal assurances and commitment in ongoing communicative actions (Gallagher and Tollefsen, 2019: 215).

As seen above, for Gallagher and Tollefsen the ‘we-narrative’ is a communicative practice, i.e. it is through the discussing of the commitment that stability is achieved. Stability is especially relevant here because it can be translated to the current
discussion as the continuation of a romantic trajectory. We-narratives need not be understood as narratives about the totality of the relationship; it can also be narrative threads that contribute to the general relationship narrative. People have shared stories about their relationship:

Intimate couples often develop we-narratives. Consider “how we met narratives.” Individuals often tell the story of how they met, either individually or jointly, and these stories often define their relationship in significant ways. Couples continue to narrate their relationship to themselves, each other, and others throughout their time together. (Tollefsen and Gallagher, 2017: 103)

Communicating about commitment and telling stories about one’s relationship certainly contributes to the stability of romantic trajectories, making the actions that continue the trajectory more likely than of this is left unspoken. It is particularly important that, when communicated, the meanings of these stories are shared by the lovers—if they are not, there will have been a disruption in mutual interpretation that can result in the dissolution of love. However, I think there is no need for explicit discussion of stories in order for the romantic trajectory to continue. After all, part of what cements a loving relationship (may that be of friendship or romantic love) goes largely unspoken.

The reason Gallagher and Tollefsen draw from communicative practices is that they are mainly concerned with the role of we-narratives in shared agency—specifically their role in the attribution of joint agency and the formation of group identity. This is an interesting topic to explore in an application of my view within the philosophy of action, but it goes beyond the scope of this thesis. I am interested here in the ‘sense-making’ feature of narratives, i.e. the way romantic trajectories are conceptualised as romantic, which makes certain actions more meaningful than others as pertaining to such trajectories.

Let us go back to the double date and imagine the aftermath from the perspective of the affectionate couple—whom I will call Diana and Aurora. Once you have got into your cab, Diana turns to Aurora and says, horrified: “They have not touched each other, not even once. These people cannot be in love!”; to which Aurora agrees. What Diana and Aurora are doing here is using their we-narrative as the blueprint to
measure whether other people are in love. Their conception of a romantic relationship includes a constant physical demonstration of affection. This is not only the way they express their love, but the way they conceive their love. I will unpack this next.

Before entering a romantic relationship, people already have an implicit conception of what being in love is. I will call this the individual conception of being in love at \( t_1 \). The lovers’ individual conceptions have different degrees of compatibility with each other at \( t_1 \). They have to be compatible enough for the people to consider they are, indeed, in love with each other, but it is rare that individual conceptions are totally compatible at \( t_1 \). For example, imagine Aurora at the beginning of the relationship, expressing her worry about Diana not being physically affectionate to her, which for Aurora is one of the defining features of being in love. “I am not sure that you are in love; you never show it”, Aurora may say. Diana, whose conception of being in love does not include a condition of physical affection at \( t_1 \), reacts to this by starting to show physical affection as an expression of her love for Aurora. So we have Diana and Aurora now both showing physical affection to each other at \( t_2 \).

Because the continued expression of physical affection forms a pattern over time, expressing physical affection becomes a component of their we-narrative or romantic trajectory at \( t_3 \). Expressing physical affection makes sense as a component of Diana and Aurora’s love for each other because it is what they have been doing in the past and what they expect to continue doing in the future. It gives stability to their we-narrative or romantic trajectory to continue engaging in the expression of physical affection. Thus, eventually, physical affection is not only an expression of their love but, also, a component of their conception of their love. Moreover, since mutual shaping is a necessary condition for love, the details of the romantic trajectory become more and more settled as the relationship advances, consolidating a couple’s we-narratives as the frame of reference to discern whether the lovers are still in love with each other.

To illustrate this claim, think of what happens when a person starts thinking that they may be falling out of love. If one day Diana found herself not being physically affectionate to Aurora, she may feel like the disillusioned philosopher in §3.3.2 feels when she finds herself bored with metaontology: Diana may start believing that she may not be in love anymore, in the same way the philosopher may start believing
that philosophy is no longer one of her values. It is tempting to say that the
expression of physical affection is an agent-independent condition for being in love,
but this would be a mistake. You and your partner may, for example, conceive that
a component of your being in love with each other is that you can sit together in the
sofa reading a book without the need of direct interaction. It may be that, come the
day when you start feeling awkward in the silence and force yourself to talk and
physically touch the other, that you may start feeling that you may be falling out of
love. What counts as being in love for Diana and Aurora does not count as being in
love for you and your partner. This is why they look at you and your partner’s
Victorian table manners and conclude that you cannot be in love: you do not fit with
their conception of being in love, which they have shaped in mutual interaction.
Relationships narratives are, then, one of the sources for their conception of what
are the necessary conditions for being in love - a conception that, in the case of
relationship narrative, is shared and shaped with the individual’s partner.

5.2.2.2. Socially shared narratives

Philosophical accounts of love tend to fall on either one of two extremes. In one
extreme, authors embark themselves on an ontological project (i.e. defining love)
without taking into account that romantic love, more than most concepts which are
involved in the shaping of our identity, cannot be understood in a sociological
vacuum (cf. Frankfurt, 2004; Nozick, 1989; Kolodny, 2003). On the other side of
the spectrum, we have the feminist accounts which, spearheaded by Simone de
Beauvoir’s *The Second Sex* (1949 [2011]), focuses on the sociological underpinnings
of romantic love and carry out their conceptual analysis in the form of a critique.

One of the most notable feminist critiques of romantic love in analytical philosophy
comes from Marilyn Friedman. According to Friedman, love entails the creation of
a shared identity which is like a “federation of identities” (1998:165) – it is then a sort
of union theory (see §1.1.1). In this federation, two separate people negotiate the
union of their individual identities, in the same way two nation-states would negotiate
the formation of a federation. This negotiation entails a great risk if one of the parties
belongs to an oppressed group – Friedman is concerned specifically with women,
given that traditional conceptualisations of love demand from women a surrender of
their autonomy. Friedman does not believe that this is a metaphysical reality inherent
to romantic relationships: “Romantic unions which nurture and affirm us can
promote our autonomy as individuals by promoting our self-knowledge, self-esteem and abilities to act effectively in concert with others (Friedman, 2003: 123–124). The main risk to women’s autonomy, according to Friedman, is not romantic love itself, but the social conceptions of romantic love, which are realised in relationships.42

Jones incorporates the feminist critique to her view. She argues that people who live in “societies structured by compulsory heterosexuality” may not even be able to conceptualise certain attitudes towards people of the same sex as being in love, given that they lack the conceptual framework (Jones, 2008: 281). Carrie Jenkins (2017) elaborates on the same idea with the following example:

[C]onsider a woman falling in love in Victorian England. The idea is that she will literally go through a different process compared to a woman falling in love in contemporary Canada. For the Victorian lady, falling in love is a matter of developing a deep and respectful (but probably rather distant) admiration for a man. Sexual desire is at best irrelevant to this process, at worst a shameful distraction. For the contemporary Canadian, however, falling in love is a matter of developing an intimate attachment that normatively includes sexual desire. If sexual desire is absent, that is at best noticeably unusual; at worst it is interpreted as showing that the feelings involved are not romantic but platonic. (Jenkins, 2017: 43, emphasis in original)

This approach is what I call a ‘purely social constructivist view’: the difference in social conceptions entails that being in love is different in kind (i.e. ‘a different process’) in virtue of which social conceptions are available. I do not endorse the purely social constructivist view (I give more details about this in §5.2.3.2). However, it is not necessary to believe that being in love is purely socially constructed to agree that there are differences between the Victorian lady and the Canadian contemporary woman. For my purposes, it is enough to accept that these are conceptual differences. That is, the Victorian and the Canadian woman have different conceptions of what being in love is, partly determined by the social conceptions available to them. This is why the sociological claim is relevant: people enter

42 Rorty takes the sociological claim even further, asserting that not only gender, but even living in a dictatorship or in an ultra-individualist and ultra-consumerist society can affect “the work of love” and “reflection over its best expression” (2016: 350).
romantic relationships with an individual conception of what being in love is (the *tl* conception) that is partly shaped by *socially shared narratives*. This approach is consistent with the view on the narrative self-concept I developed in chapter 3.

In my discussion about the narrative self-concept, I have shown that Minimal Narrativism departs from the Strong Narrativist requirement for story-telling, i.e. personal narratives having to adjust to a genre like fictional narratives do (§2.3.3; §3.4.2). Anti-narrativists rightly object to this interpretation of story-telling. However, I called for a re-interpretation of story-telling as narrative models: “the story or stories that we would want told about who we are” (Hutto, 2016: 26). I do not think that there is any room for objection to the fact that we have numerous narrative models about romantic love available in fiction, for example. From the courtly love of Tristan and Isolde to Shakespeare’s romantic tragedies, all the way to Charles Bukowski’s love poems, Joni Mitchell’s songs and *The Diary of Bridget Jones* or Rachel and Ross from *Friends*, fictional stories about romantic love reflect different expressions of love that make their way into people’s conceptualisation of is to be in love. They do so by becoming social scripts: as I explained in §3.4.2, some of these narrative models are what Lindemann calls master narratives, socially shared narratives which contribute to the shaping of certain experiences. The Victorian woman has a social script available as to what being in love is, given that she lives in a society where respect and deference are components of true romantic love while passion and sexual desire are mere signs of fallible infatuation. The Canadian woman has a social script with almost opposite content.

Different cultures provide with different social scripts, but also different sub-groups within a culture: an American who is raised a Mormon in Salt Lake City has a very different social script for romantic love from a New Yorker, and person that grows up in a *kibbutz* has a very different social script for romantic love than a London-born Orthodox Jew. At *tl*, individuals arrive at the relationship with a conception of romantic love that is partly shaped by master narratives which will feed into the relationship we-narrative that is subsequently formed. Again, this conception will provide an individual with certain conditions for an instance of personal love to count as being in love.
5.2.2.3. Individual narratives

Finally, the individual self-concept directly influences the conception that a person has of what being in love is at \( t \). This is an obvious claim that can be taken as a given: if socially shared narratives and relationship narratives shape one’s conception of what being in love is, it is easy to see how individual narratives (where one has lived, what one has done in one’s life, one’s defining traits) equally contribute to this conception. I want to highlight here a different, more illuminating way in which the individual self-concept contributes to a person’s conceptualisation of what being in love is.

Rorty has an example which leads to the claim I make here. She gives an example where Louis, who used to be in love with Ella (but is not anymore), starts a relationship with a new person, Gloria: “Presumably Gloria does not want to inherit Louis’s love for Ella: she wants Louis to love her in a wholly different way, defined by the two of them” (Rorty, 1987: 402). With this, Rorty wants to highlight the active role of the lovers in mutual shaping: Louis cannot just reproduce what he did with Ella in his relationship with Gloria, since he must be dynamically permeable to Gloria, who is a different person. However, later Rorty says that after loving Ella, “Louis’s subsequent history, his new loves, joys, indignations, the details of his continuing evaluation – even his love for Gloria – are affected by his loving interaction with Ella” (Rorty, 1987: 404). I am going to go further and say that Louis’s love for Gloria is not only affected by, but partly constituted by his love for Ella.

In §3.4.1 I showed that in my Minimal Narrativist view the self-concept is diachronic, even without requiring diachronic self-experience in the way portrayed by Strawson (i.e. consciously experiencing oneself as a persisting being). I explained that present characteristics are partially constituted by past characteristics. Whether a person endorses, repudiates or is alienated from her past characteristics shapes the way she understands herself now. Here, I add that the trait of being in love is also diachronic in that way.

It is not uncommon to be in more than one romantic relationship throughout one’s life.\(^{43}\) Mutual shaping changes an individual’s self-concept in ways that sometimes

\(^{43}\) As an illustrative datum, the global rate of divorce has more than doubled between 1970 and 2008 (Wang and Schofer, 2018).
survive the relationship—both Rorty above and Cocking and Kennett acknowledge this. Emma’s acquired interest in hiking may continue after breaking up with Seb, for example. As I have explained above, relationship narratives contribute to an individual’s conceptualisation of what being in love is. A person who has several romantic relationships over time will then subsequently develop different conceptions of what being in love is.

Imagine that Diana and Aurora were to terminate their relationship, where the conceptualisation of being in love included being physically affectionate. Looking back, Diana may have different attitudes towards the relationship narrative she shared with Aurora. She may endorse it, thinking that her relationship with Aurora showed her that being affectionate is indeed constitutive of being in love. She may repudiate it, thinking that being affectionate put her in an uncomfortable position constantly and led to the diminishing of her autonomy. She may even feel alienated from it, not recognizing herself in that relationship narrative, in the same way the serious matron did not recognize herself in the party girl.

In line with my view that being in love is a trait of the self-concept, and that traits are partly constituted by one’s relation with past traits, I believe that this relation with past experiences of being in love partly constitute one’s present experience of being in love. The relation that Diana has to this past relationship narrative with Aurora will be a fundamental component of her new relationship with Julio. If she endorses it, Diana’s conception of what being in love is at \( t_1 \) in her relationship narrative with Julio will include being physically affectionate. If she repudiates it, Diana’s conception of what being in love is at \( t_1 \) will not only not include being physically affectionate: were Julio to demand physical attention, that may be a potential source of conflict between them. What I do here is reject Rorty’s initial claim about Louis’s love for Gloria: Julio cannot, as Rorty initially says, require from Diana that she loves him in a ‘wholly different way’ to the way she loved Aurora. Diana’s love for Julio is constituted by her love for Aurora: it is part of Diana’s conception of what being in love is at \( t_1 \). In that sense, an individual’s conception of being in love is partly shaped by that individual’s narrative as a lover.
5.2.3. A relatively agent-dependent view

I have now filled the gap that both Jones and Helm left unexplained. People conceptualise what being in love is through three levels of narrative: the relationship narratives, the socially shared narratives and the individual narratives of the lovers. This conceptualisation is what provides people with necessary conditions for categorising their personal love as an instance of being in love. For some people, this includes a drive to demonstrate affection physically; for others, the ability to sit down together in silence. I will insist, it is not just a difference in expression of being in love, but as what counts as being in love: recall my examples in §5.2.2.1 and my comparison with the disillusioned philosopher. What being in love is, then, is agent-dependent. Does this mean that there is, across different times, places, and cultures, one unique conception of what being in love is for each existing individual? I do not think it does, so I am going to qualify the claim on the agent-dependent nature of being in love.

The truth is that there is bound to be a great overlap between individuals, for two reasons. The first one is that socially shared narratives are, as the term indicates, shared by a considerable number of individuals. Many Victorians will have had the same conception of what being in love is as the Victorian lady in the example. This is even more so in the case of the Canadian woman, who lives in a globalised world where socially shared narratives have an outreach unprecedented in human history. So, although relationship narratives and individual narratives will have a greater degree of difference, people living in the same society will already have a lot in common at $t_1$ due to the socially shared narratives available to them.

The second reason is that it cannot be ignored that romantic attachment has a biological basis and at least correlates with certain neural and hormonal processes. Since human individuals are of the same species, this biological basis already lays common ground between individuals at $t_1$. Although I do not believe that biology by itself can explain what being in love is, it would be unwise to ignore this fact. Discussing the biological basis of love would significantly deviate the discussion.
here. I merely accept the conventionally accepted fact that love (or at least something in its vicinity) has such a biological basis.44

Hence, I claim that being in love is ‘relatively agent-dependent’: individual conceptions of being in love vary in virtue of people’s individual, relationship and socially shared narratives, but there is an overlap between large groups of people given that they live in the same society and they belong to the same species. It is not surprising, then, that for a large majority of people in Western society, sexual desire (which is also a biological drive) is a condition for an instance of personal love to count as being in love. This approach allows me to qualify the claim that every person’s conception of being in love is completely unique, a claim which is admittedly bizarre and counter-intuitive.

Nevertheless, the situation at the double date is explained by the room for agent-dependency that my view makes. When Diana and Aurora exclaim that you and your partner cannot possibly be in love with each other, they are not committing a conceptual error. This is because, if it was them in your situation, that would mean that they are not in love with each other, or at least cast suspicion on it. You and your partner are of the same species and live in the same society than Diana and Aurora, but given the unique nature of your relationships and your individual narratives, you have conceptualised what being in love is (i.e. which are the necessary conditions for being in love) in a different way.

This leads back to the other challenge I posited in this section. Is there a difference between romantic love and friendship? The answer, as above, is not metaphysically (i.e., not in kind), but different individuals may conceptualise them as different for them without incurring in conceptual error, given that this is a relatively agent-dependent conception. In order to further support this claim, I offer some examples which back the diffusion of ontological limits between romantic love and friendship, while separating my theory from purely social constructivism.

5.2.3.1. Friends and lovers in history

The agent-dependent component of my view may still find some resistance given the above observations on how, in everyday language, there seem to be objective

differences between romantic love and friendship. To completely allay worries on this issue, I offer two examples.

Firstly, I bring in Michel de Montaigne’s study on friendship. For Montaigne, friendship is the highest form of love. The following are his observations on the moment he met his friend Étienne de La Boétie:

We were seeking each other before we set eyes on each other—both because of the reports we each had heard, which made a more violent assault on our emotions than was reasonable from what they had said, and, I believe, because of some decree of Heaven: we embraced each other by repute, and, at our first meeting, which chanced to be a great crowded town-festival, we discovered ourselves to be so seized by each other, so known to each other and so bond together than from then on none was so close as each was to the other (Montaigne, 1580 [2004]: 10).

Secondly, Simon May’s historical study of the philosophy of love describes as follows the conception of eros in Ancient Greece, as reflected in Plato’s Symposium:

Most of the speakers [in the Symposium] take it for granted that only homosexual love between a man and a youth is conducive to the noblest aspirations, such as the pursuit of learning, politics and refined manners… By contrast, some of the speakers… regard heterosexual love as altogether inferior: a largely bodily impulse whose chief value is physical procreation. Love between a man and a woman, these speakers hold, is no starting point for spiritual growth. To the point where, if an older man were to desire a younger woman in this way—in other words for sexual intimacy leading to spiritual friendship and pregnancy of the soul—they might both need to pretend that she is a boy, or at least ignore the fact that she isn’t. (May, 2011: 40–41)

I said at the beginning of this section that, in everyday language, people seem to differentiate between romantic love and friendship, which may give grounds to their differentiation in kind. However, if we put these quotes to the test of everyday language, I suspect they would hardly be considered examples of friendship (in the
case of Montaigne) or reflecting acceptable conditions for romantic love (in the case of the Ancient Greeks).

To the contemporary eye, Montaigne plausibly seems to be describing someone he is in love with, not a friend. Montaigne shows how a widely shared conception of what being in love is (such as defining meeting for the first time as ‘some secret appointment of heaven’) may lead to a mistaken categorisation of his experience. Montaigne thinks of La Boétie as a friend. He loves La Boétie with the highest form of love, which for him is friendship, not romantic love.

Equally, the Ancient Greeks as portrayed in May’s analysis have a description of *eros* that would now be considered at the same time highly moralised (having the noble as its aim) and highly immoral (requiring from women to pretend to be boys in romantic relationships). This is precisely because the Ancient Greeks’ socially shared narratives friendship and romantic love were socially constructed, parts of which are at odds with contemporary socially shared narratives.

The reason why Montaigne’s friendship and the Ancient Greeks’ romantic love do not immediately look to the contemporary eye as genuine cases of friendship or romantic love is, *precisely*, because we are looking at them with a different conception of what being a friend at what being in love is. The reason for this discrepancy is that our conceptions are relatively agent-dependent, so that we ought not to extract metaphysical conclusions from the everyday distinction in this case. A persistent objector would have to argue that Montaigne was actually in love with La Boétie, or that the Ancient Greek had a morally wrong conception of *eros*. The latter objection, being a moral one, is outside the scope of my argument. As for the former, the burden is on said persistent objector to find more convincing arguments for that position than Montaigne’s own personal report; a strategy that, I would say, seems highly patronising to the philosopher.45

———

45 Montaigne’s description of his feelings for La Boétie may have started to become understandable as friendship in Westernized societies with the progressive acceptance of the idea of ‘bromance’. This term, which is a portmanteau of ‘brother’ and ‘romance’, is “an exceptionally tight, affectional, homosocial male bonding relationship exceeding that of usual friendship, [which] is distinguished by a particularly high level of emotional intimacy” (Bromance, n.d.). I believe that this type of intense friendship is best categorised as ‘romantic friendship’, which can happen between any combination of genders. The fact that some people would differ in characterising Montaigne’s relationship as being in love or having a romantic friend does, in fact, offer further support to my claim of agent-dependency.
5.2.3.2. Off-the-mainstream love

In light of the above examples, it may then seem that the relatively agent-dependent claim is superfluous. It may seem more theoretically economical to say that romantic love and friendship is socially constructed, as some feminist philosophers do with other emotional phenomena (see fn. 30). However, a pure social constructivist view on being in love would ignore the biological basis of love and, more importantly, would leave out off-the-mainstream conceptualisations of love. I focus on the second oversight, which is the most relevant for my view.

As seen above, Jones claims that living in a society where heterosexual romantic love is the norm or even the only legal form of romantic love (which is how I interpret her ‘compulsory heterosexuality’ observation) may not allow queer people to conceptualise their experience as being in love. Jenkins has a similar example, where she describes a lesbian couple in late-nineteenth-century England. The women in this couple, Jenkins says, may have all the biological manifestations of being in love but, nevertheless, not be able to engage in the type of activities considered romantic, like marrying each other, expressing physical affection or parenting together (Jenkins, 2017: 108). According to Jenkins, a purely social constructivist view would establish that these women could not be in love due to the lack of socially shared narratives to categorise their experience in that way.46

This position is untenable, for several reasons. Firstly, it is factually inaccurate. There were lesbians in late-nineteenth-century England. One prominent example is Anne Lister, best known by her nickname ‘Gentleman Jack’. It was also in the Victorian era when Oscar Wilde was tried for his relationship with Lord Alfred Douglas. In 1682, the marriage of singer Arabella Hunt was annulled upon public discovery that her ‘husband’ was, in fact, a cisgender or biological woman. Earlier than that, several English kings, such James I (sixteenth/seventeenth century) had documented

46 For a historical-psychological defence of a purely social constructivist view on romantic love, see Anne Beall and Robert Sternberg (1995). Beall and Sternberg’s view has some point in commons with my view, mainly that socially shared concepts and the actions in a specific relationship contribute to define what being in love is. Stenberg (1995) even notes the influence of narrative in creating such conception, as I do here. His table on types of love stories (Sternberg, 1995: 143) is particularly illuminating as to how we-narratives may deviate from universal or moral standards (for example he includes narratives of ‘Horror’, ‘Addiction’ or ‘War’ as common relationship narratives). However, their view is that love is wholly socially constructed, a claim that I reject in this section.
homosexual relationships. These are some hand-picked examples in the history of queer love in Britain alone.

Secondly, as these examples show, a purely social constructivist view ignores that people’s conceptualisations of romantic love may go against or depart significantly from socially shared narratives. Otherwise, it is hard to explain these examples of queer love in overwhelmingly heterosexual societies. My claim here is that compulsory or generalised heterosexuality does make it more difficult for people to conceptualise certain experiences of romantic love, as I have said in §5.2.2.2; but it does not make it impossible. One way of explaining this is on biological terms: people involved in these off-the-mainstream romantic relationships just had stronger physical drives, or a stronger will to act in accordance with these drives, than most people in their situation. However, I think this explanation is insufficient and makes poor justice to queer experiences of love. I have defended in this thesis that being in love is an expression of one’s life worth living. I said in §2.1.3 that this is both a matter of discovery and invention: it is an exercise of a person’s autonomy. This claim loses its force all that happens is that ‘biology pushes through’, so I think it is unfair to use that argument in the case of queer experiences (especially if it is used in the case of queer experiences only). Explaining these off-the-mainstream relationships as a result of agent-dependency of the necessary conditions for being in love is, as I will show next, a better-suited explanation than a biological one.

I bring back once more the example of the double date. The two couples have differing conceptions of being in love, which leads them to think that the other couple either is not in love or has a defective way of showing it. The advantage of the relatively agent-dependent approach is that it allows for both couples to be considered instances of being in love, without either of them necessarily incurring in conceptual error. In other words, my theory can explain odd instances of being in love as plausibly such. In the case of the double date, these couples were odd for each other, but a reader may find the two experiences as plausible instances of being in love. Whether my claim on agent-dependency is convincing partly rests on the acceptance of this. So now, I am going to put forward an example which seems a prima facie odd instance of being in love, or not an instance of being in love at all. I do so to show how an agent-dependent view can explain off-the-mainstream
instances of being in love in a way that a biological approach cannot. I draw from Paul Thomas Anderson's film *Phantom Thread* (2017).

The main characters of the film, Reynolds and Alma, are involved in a relationship where the only way for them to be together involves a bizarre cycle. First, Alma poisons work-obsessed Reynolds’s soup, which puts him at death’s door. Then, Alma nurses him back to health, which rekindles Reynolds's love for Alma until they fight again and re-start the cycle. Reynolds knows this and participates in the cycle voluntarily: his narcissism and self-obsession mean that he can only be vulnerable and fully love Alma and let her love him if he is completely defenceless. This looks like an unhealthy type of relationship, a suspicion that is exacerbated by the way Reynolds treats Alma throughout most of the film – more like an object that he owns than like a person in her own right. I do not intend to claim that Reynolds’s surrender to Alma in their bizarre cycle condones all the times that he is plainly abusive to her. Even considering the poisoning-nursing to health cycle in isolation this relationship seems *prima facie* as imbalanced and morally problematic. However, it is not impossible that Reynolds and Alma are in love with each other: it is just that their conception of being in love, and the expression of that conception, deviates considerably from socially shared narratives. Their bizarre cycle is their relationship narrative, and hence is constitutive of their conceptualisation of their being in love. As per the trait view, it is because Reynolds realizes Alma is part of what makes his life worth living, and that his character flaws prevent him from being able to live according to such conception unless he is totally defenceless, that he willingly agrees to the dynamic (“I’m getting hungry”, he tells Alma at the end of the film).

The point I am trying to make here is that in the *Phantom Thread* story there is no clear biological drive that warrants such an off-the-mainstream relationship narrative. I do not think that this narrative warrants the label of socially shared; quite on the contrary, it seems to be against most socially shared narratives of what being in love is. With this extreme example, I aim to illuminate how a relatively agent-dependent view on the conceptualisation of being in love may explain even the oddest cases as plausible instances of being in love in a way that a pure social
constructivist view, or a purely biological response to social constructivism, cannot. My view is not that Reynold and Alma are in love. It is, instead, that given that the conditions for being in love are agent-dependent, it is plausible that they are. Any other approach entails an imposition of conditions for being in love (biological, social, or morally based) which I am rejecting here, given that my aim is to offer a definition of being in love which obtains universally. The only option for such a universal definition is that being in love is (relatively) agent-dependent.

Admittedly, the fact that a seemingly abusive relationship like Reynolds and Alma’s can plausibly be an instance of being in love is, to borrow Kennet’s expression on morally dangerous friendships, deeply uncomfortable (see §4.1.3). However, this approach rightly captures the risks that come with being in love. Recall Edward Harcourt’s observation that love can be ‘the unmaking of lives’ (§1.3.3). Although Harcourt’s observation is made from a different approach to mine, this risk is fully incorporated in my view. The risk entailed by being in love is, crucially, the main focus of the feminist critique of romantic love which I introduced above (§5.2.2.2). De Beauvoir says that, for women, love is chosen “enslavement”, a harmful obliteration of their autonomy that “becomes a religion” for them (1949 [2011]: 700). Friedman provides with a long list of ways how power imbalances on the basis of gender can damage a woman’s autonomy in romantic relationships, by making her submerge her own identity into the loved person’s (2003: 125). What feminist critics highlight is how romantic love can be dangerous for autonomy, and my example illustrates this danger. But, in the end, it is romantic love that they are criticising: my aim here is to show that off-the-mainstream instances, even damaging ones, can be plausibly instances of romantic love, and that they can be explained as such only in the context of a relatively agent-dependent view.

Having justified my position on the fact that conditions for being in love are relatively agent-dependent, I have now offered full support for the trait view. Being in love is a trait of the self-concept, which articulates a person’s understanding of herself and the world; an understanding that is acquired through mutual shaping

---

47 It should be clear that the only parallelism I am drawing between queer romantic love in heteronormative societies and the Phantom Thread relationship is that both are beyond socially shared narratives of what being in love is, and nothing else. The fact that queer love is not a shared narrative in a given society makes that society morally defective, while that is probably not the case in a society that does not share the Phantom Thread characters’ conception.
embedded in a romantic trajectory. That mutual shaping is the only necessary condition for an experience to count as an instance of being in love: apart from mutual shaping, these conditions are relatively agent-dependent. In that way, I have concluded the articulation of my narrative account on being in love, which here I have called the trait view. Finally, I briefly deal with an objection that stems from the fact that the trait view rests on a narrative theory of the self-concept.

5.3. Can Galen Strawson fall in love?

In §2.1.3 I stated that narrative theory was a promising framework to define being in love as a trait of the self-concept, given its focus on the relationality and the historicity of the self-concept. In the last two chapters, the turn to narrative theory has been fully justified. The Minimal Narrativist view I developed in chapters 2 and 3 has provided with the tools to explain mutual shaping, trajectory-dependence and being in love as a defining trait, as well as the three-tiered conceptualisation of being in love.

Now, this view may raise a worry with regard to non-narratives, i.e., people who do not formulate narrative products to understand themselves. In §2.3.2 I accepted that this is a valid form of self-understanding which is accommodated by Minimal Narrativism. However, given the entrenchment of the trait view with Minimal Narrativism, a question may arise: does this view entail that non-narratives cannot be in love, or that their love is somehow defective? Strawson conveniently anticipates this very question from narrativists:

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, 2007: 109)

The first answer to Strawson is, like I showed in §2.3.1, that here he is portraying a Strong Narrativist view, which—arguably— requires explicit recollection of the past and diachronic self-experience in the sense of experiencing oneself as a persistent
being. I have already shown that diachronicity is not understood in these terms in my Minimal Narrativist view on the self-concept. The diachronicity of the self-concept is, instead, provided by the fact that past characteristics are constitutive of present characteristics via the relation an individual has to them.

The second answer is that Strawson misconstrues what a narrative view on being in love requires. It seems that Strawson is implicitly pointing in the direction of constant, detailed or precisely accurate explicit recollection, that is, telling actual stories about the shared past. That, which is a charitable reading of Strawson’s claim, is simply not required in the trait view (see §5.2.2.2 on we narratives). Otherwise, an uncharitable reading of Strawson’s claim strikes as a case of someone who is not in love. If we are to take seriously that one need not remember or value one’s shared past with a partner, imagine sitting down with your partner and bringing up some moments of your shared life. “Do you remember when we met at University?” “No.” “Do you remember that bar when we used to go?” “No.” “Do you remember our wedding? The birth of our first child?” “Whose wedding?” The partner in this example has so little receptivity to mutual shaping that it has not even shaped their memories. So, unless they have a severe memory problem or can quote other shared experiences that have made some impact on them, this person does not fulfil the minimum requirement to consider that they are in love. In any case, I choose to do the most charitable reading of Strawson’s claim and simply reject that telling fully-fledged stories about the shared past is a requirement of my view.

What is, then, my answer on whether non-narratives can be in love? The answer is that non-narratives can, of course, be in love. The relationship narratives that I have described in §5.2.2.2 are not always articulated, and when they are, they are not always articulated in the shape of fully-formed stories. The same goes for socially shared and individual narratives. If Strawson’s idea of describing a relationship is “We met, we fell in love because we liked each other, and we are still together”, there is nothing that invalidates that experience as a genuine experience of being in love – as I said, this is a relatively agent-dependent matter.

The Minimal Narrativism view on the self-concept which the trait view is built upon only requires narratives for maximal understanding. Narratives, I argued in §3.1.2, provide with maximal understanding given that they have higher explanatory power than causal explanations: unlike causal explanations, narratives can explain what
people feel and judge about certain situations. By looking at the narratives a person formulates, we can also access what they find salient or worthy of interpretation, which is articulated by their self-concept. What a non-narrative will never achieve is maximal understanding. Without engaging in narrative thinking, one cannot obtain explanations as to how one felt and how one judges the different events in a relationship, which led to the current we-narrative. Neither can one identify the socially shared narratives or the events in one’s individual narrative that have partly shaped one’s conception of what being in love is. A non-narrative will also struggle to explain the way their former loves have shaped their current love.

That, however, does not make non-narrative love inherently defective. The truth is that many people—not only non-narratives—do not have maximal understanding of what is for them to be in love. That is in line with my view of both harmonic and non-harmonic love. I have defined what being in love is, but I accept that a full description of one’s individual conception will remain mostly inaccessible for the majority of people. The trait view, I shall recall, is not a theory of reasons for love. Within my narrative view on being in love, it is still perfectly fine to answer like Montaigne did to the question of why he loved La Boétie: “If you press me to say why I loved him, I feel that it cannot be expressed except by replying: ‘Because it was him: because it was me’” (1580 [2004]: 10).

Conclusion

In this chapter, I have developed and offered support for a novel theory in the philosophy of love: the trait view, according to which being in love is a trait of the self-concept. I have shown how the trait view can incorporate the right claims of caring view, which was the starting point of this thesis. I have also fully incorporated mutual shaping, which is a necessary condition in my account, and trajectory dependence, which is a justificatory reason for mutual shaping. Throughout this thesis, my aim has been not so much to prove that these views are wrong, but to show that there is a better way to arrive at the correct claims they make without over-idealising love or accepting that love justifies itself.

The trait view and the narrative view of the self-concept I have developed here open a series of questions that can be answered within the account. Mainly, although I have argued that there are no agent-independent necessary conditions for being in
love (besides from mutual shaping), we can still ask which are the reasons to love a particular individual. I think this thesis opens a route to understand the rationality question through some sort of ‘narrative fit’ between the lovers. Also, we may want to find out which are the best and the worst conceptions of being in love. Defining this will rest, I believe, in the relation between being in love, the self-concept and autonomy. These are just some sketched future applications of the project in hand which directly related to the main questions (rationality and autonomy) in the literature.

The trait view explains what being in love is for the lover while at the same time fully accounting for the importance of the loved person’s identity. The loved person is part of the lover's conception of a life worth living, a co-protagonist and co-creator of the story the lover wants to be told about herself. With being in love as a defining trait, the loved person contributes to shaping the lover’s understanding of herself and the world. In the trait view, being in love is not a process where the lovers appraise each other’s value. It is a process of becoming together. Crucially, being in love can improve the lover’s life as much as it can destroy it. For some people, being in love is a process of enlightenment; for others it is a process of self-destruction. My aim has been to incorporate everyday experience of being in love as well as the basic claims of the feminist critique without my analysis taking the form of a critique.

In the trait view, being in love is individually conceptualised, on the basis of the life one has led, the lovers one has had, the love relationships one has been involved in, and the society one lives in. For that reason, being in love is agent-dependent; relatively so, given that there is bound to overlap between individuals living in the same society. Thanks to the condition of agent-dependency, the trait view can explain why off-the-mainstream relationship narratives can be plausibly considered instances of being in love. A critic may require a more precise articulation of the necessary conditions for being in love, but to that critic, I would say that such an approach is by itself a methodological mistake. Trying to define being in love through a list of necessary and sufficient conditions is already, in my opinion, a misrepresentation of the phenomenon.

The fact that the trait view rests on a fully supported and plausible theory of the narrative self-concept means that it can incorporate all the advantages of the narrative theory without giving room for anti-narrativists to reject it tout court. The
fact that this is a narrative account does not allow for an immediate anti-narrativist dismissal of the trait view: first, they would have to reject the solutions I have offered to the challenges they present.

In summary, by departing from traditional approaches, the trait view introduces new issues on the debate of love (narrativity), develops issues that have only recently been incorporated to the mainstream discussion (agent-dependency and social constructivism) and offers an alternative route to understand widespread positions (caring and identity). In the trait view, both the lover and the loved person are worthy objects of analysis. I believe this is a better way to capture what being in love is: not only the way that one cares about the loved person, but the way the loved person becomes, for one, a new way of understanding one’s life.
Bibliography


Rorty, A.O. (1987). The historicity of psychological attitudes: Love is not love which alters not when its alteration finds. Midwest Studies in Philosophy


