

On Why Emotional Recalcitrance is Only Apparent

MPhil Stud

Cristina Diana Craciun

University College London

2021

I, Cristina Diana Craciun, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Cristina Diana Craciun, 30th of August 2021

Abstract

In this thesis I address the topic of emotional recalcitrance, the phenomenon in which one's emotions are in a state of conflict with one's judgment. It might be assumed that this conflict indicates that such emotions must be irrational, since they go against our judgment. However, a plurality of voices in the literature have tried to argue that these emotions are not irrational, contrary to appearances. For instance, some emphasise the underlying biological processes of these emotions, leading to views on which recalcitrance is potentially a-rational. Another strategy is to focus on the differences between how judgments and emotions relate to their content, for instance by arguing that emotions are more like perceptions, which are not irrational even when conflicting with judgments. I will argue against both types of responses.

On my view, recalcitrant emotions are not irrational because they are not in conflict with the judgments we take them to be in conflict with. That is, subjects misunderstand the content of their own emotion and compare it with the wrong judgments – the emotion, however, is therefore not in any breach of rationality. This assumed conflict is the source of irrationality, rather than the emotion itself. Emotional recalcitrance is, then, the persistence of a misunderstood emotion.

I will organise this work as follows. In the first chapter, I will offer a taxonomy of existing theories of emotions and assess their approach to the challenge of recalcitrance. In the second chapter, I will offer a new lens for discussing recalcitrance: a distinction between *strong* and *weak* recalcitrance. I will show that my account is the most promising way of accounting for this distinction, while also being consistent with the majority of existing views of emotions. In the final chapter, I address potential issues for my account, particularly issues concerning normative criteria.

Impact statement

This thesis provides an original response to the problem of emotional recalcitrance, which will be impactful both for future research in the philosophy of emotions and for psychology and therapy. First, this work offers a solution to the challenge of emotional recalcitrance that is independent of any one theory of emotions, thereby contributing to the overall mission of developing a satisfactory account of the emotions. My account is particularly useful to judgmentalism, since it allows it to avoid the implausible implications recalcitrance would have on it. Furthermore, I offer a detailed analysis of the concept of intentional object and develop a new way of understanding it, which will meaningfully contribute to further theoretical discussions about the emotions.

Secondly, my conclusion is relevant for psychology and therapeutic practices because it encourages better understanding of the content of an emotion, with the implication that a thorough understanding of these objects will lead to resolution of these problematic emotions. In particular, my account offers solutions for dealing with emotions that resist our judgments, which often is the case with trauma and other compulsive emotions. My account favours a type of self-understanding as a precursory to emotional healing, which is consistent with current empirical approaches but recontextualises them.

Contents

Introduction	5
Chapter 1: What are the emotions and what is recalcitrance?	8
1. The Philosophy of Emotions	8
2.1 Feeling theories	9
2.2 Advantages and disadvantages of feeling views	10
3.1 Judgmentalism or evaluative judgment theories	14
3.2 Advantages and disadvantages of judgmentalist or evaluative accounts	19
4. Perceptualism	22
a) Strong perceptualism	23
b) Weak perceptualism	26
5. Attitudinal and Motivational views	29
a) The Attitudinal View	30
b) The Motivational View	34
6. Recalcitrance	38
Chapter 2: Does (the conflict of) emotional recalcitrance exist? The challenge from strong and weak recalcitrance	43
1.1 Setting the stage	43
1.2 Relevant terminology and key concepts	47
2. Two different solutions	51
2.1 The problem is the emotion	51
2.2 The problem is the relation between the emotion and judgment	54
3. My solution	60
4. Concluding thoughts	64
Chapter 3: The theory defended	67
1. The worry of introspection	69
2. Intermezzo: Emotions about possibilities?	70
3. Emotional justification and the worry of continuous emotions	73
a) Cognitive rationality	73
b) Practical rationality	77
4. Conclusion	80
Conclusion	81
Bibliography	83

Introduction

The puzzle of recalcitrance has been central to the inquiry into the nature of emotions. Briefly, an emotion is recalcitrant if it conflicts with a judgment, for instance fear of a spider one knows to be harmless.¹ This is a very common phenomenon, as we often find ourselves puzzling over an emotion that would not go away even though (we think) we know better. One important characteristic of these episodes is that we regard these emotions as irrational in virtue of their conflict with judgments. Yet why would this common phenomenon be so central to the definitional question of emotions? One answer is that many views of the emotions, especially early ones such as judgmentalism and feeling theories, struggle to account for it.

First, judgmentalism, which argues that emotions are a type of judgment,² would lead to a very strong conclusion: that we are contradicting ourselves. If our emotion is a judgment that p , and it conflicts with a belief that $\neg p$, then we are contradicting ourselves and we are therefore fully irrational. Of course, this conclusion would be too strong – recalcitrant emotions may have irrational aspects, but they are not as irrational as full-blown contradictions between two judgments. In other words, consciously thinking to oneself both p and $\neg p$ seems more irrational than fearing a spider while thinking it is not dangerous. We are, in other words, not *that* irrational.³

Second, feeling views, which reduce emotions to their phenomenal character (how they feel to the subject),⁴ struggle to explain why recalcitrant emotions would be even remotely irrational. Since they are mere feelings such as pain, tiredness, or feeling ticklish, they simply cannot stand in conflict with beliefs. We cannot have irrational pain, or irrational tiredness, so we cannot have irrational fear either. Feeling theories thus have the opposite problem from judgmentalism: they make these emotions not irrational enough.

¹ Majeed (2020), Grzankowski (2017, 2020), D'Arms and Jacobson (2003, p.129).

² See chapter 1, section 3.1 for a more in-depth analysis of Judgmentalism.

³ Additionally, as Dow (2009) puts it, judgmentalism would render such situations unusual or exceptional, since we do not normally contradict ourselves so blatantly. However, since recalcitrance is a fairly common occurrence, it should not feel exceptional, which is therefore an issue for judgmentalism (p.146).

⁴ See chapter 1, section 2.1 for a more in-depth analysis of Feeling theories.

Thus, the puzzle of emotional recalcitrance has become a test case for any view of emotions. Theories of emotions must therefore find a way to balance i) the fact that recalcitrant emotions meaningfully conflict with judgment and thus seem to exhibit irrationality and ii) they are not as irrational as a judgment conflicting with another judgment.⁵

This puzzle is consequently not only useful in identifying a promising theory of the emotions, but it also opens the door to discussing the connection between reason and emotions, states which have long been thought to stand in opposition to each other.⁶ But if emotions were thoroughly irrational, arguably every single emotion would feel like recalcitrance – yet they do not. When one grieves the loss of a loved one, when one celebrates a victory, when one is afraid of a shark, one's emotion seems to be doing something right, it seems appropriate to the situation. Likewise, fear of a towel, anger at somebody who did nothing wrong, joy over a tragedy, all seem unjustified – they fail some normative criteria. Recalcitrance is therefore a great starting point for discussing the normativity of emotions.

Taking all of these ideas into account, this work will be structured as follows. In the first chapter I offer a general survey of the different views of the emotions discussed in the literature. I discuss feeling views, judgmentalism, perceptualism, attitudinal and motivational views, and then I focus on how each of them addresses the challenge of recalcitrance.

In the second chapter, I propose a new way of discussing recalcitrance. I start by noting an ambiguity in how recalcitrance is defined. Since it is defined simply as a conflict between an emotion and a judgment, it also includes emotion states that may be in conflict with existing judgments of which we are unaware. These emotions importantly change as a result of that judgment, particularly once we become aware that we do hold that judgment. This is what I call 'weak recalcitrance', and by including it in the discussion of recalcitrance we fail to do justice to an essential feature

⁵ See Dow (2009, pp.146-147) for a more detailed outline of the criteria.

⁶ As Solomon (1977, p.35) puts it, 'The myth of the passions', according to which emotions are irrational forces beyond our control, has dominated most of the history of philosophy, yet has recently been challenged both within philosophy and the empirical sciences.

of recalcitrance, namely evidence-unresponsiveness. What characterises recalcitrance, I argue, is its *persistence* in the face of contrary evidence (I will call this persisting conflict strong recalcitrance).

Once I establish this distinction, explaining the phenomenon of recalcitrance becomes explaining how it is that some emotions are evidence-unresponsive. To that aim, I challenge the very idea that there is a conflict between beliefs and emotions, with the goal of showing that these emotions are only evidence-unresponsive because they are faced with the wrong evidence. That is, we misunderstand the content of our emotion, and thus assume it is in conflict with the wrong judgment. The conflict, then, does not exist at all – thus, the emotion is not actually irrational. However, it seems to us like it is irrational simply because we assume it is in conflict with a certain judgment. This account thus satisfies both i) and ii).

I dedicate the third chapter to defending this account from potential worries, specifically worries concerning the normativity of recalcitrant emotions as I describe them. In short, I argue that these emotions are about more than we give them credit for: they are not only about, say, lipstick traces on one's partner's shirt. Instead, they are about any number of possibilities, implications, explanations, or general mental states related to those lipstick traces. For instance, they could be about who they may be cheating with, why, what that means about your relationship and their love for you (or lack thereof).

Since I argue that recalcitrant emotions are about such complex mental states, beliefs such as 'it is not lipstick, it actually is tomato sauce' may fail to relieve the anger simply because the anger is about how they could potentially be seeing A behind your back or how they may not love you anymore, which remain possible regardless of whether the redness is due to tomato sauce or not. Thus, the emotion seems to still be justified given its object: who would not be angry at the prospect of being cheated on? Or afraid at the possibility that a gorilla escapes and attacks you? Of course, there may be problems with such emotions about possibilities, so I will spend the third chapter discussing possible normative criteria for the emotions and inquire whether emotions about possibilities do satisfy them.

Chapter 1: What are the emotions and what is recalcitrance?

1. The Philosophy of Emotions

The puzzle of recalcitrance has greatly influenced the philosophical discussion of the emotions: any account of them must be able to provide an explanation of this all-too-common phenomenon. Thus, I will begin by outlining the current landscape of the philosophy of emotions, which will provide us with the necessary background to fully grasp the weight of the challenge of recalcitrance.

To that aim, I will first spend the rest of this section outlining some key principles and caveats concerning the philosophy of emotions. Then, in section 2.1, I introduce feeling views, followed by the advantages and disadvantages of these views in 2.2. Following the same structure, I introduce judgmentalist views in 3.1 and their advantages and disadvantages in 3.2. In section 4, I introduce perceptualist views, first discussing strong perceptualism in 4.a) and weak perceptualism in 4.b). In section 5, I introduce motivational and attitudinal views, addressing attitudinal views in a) and motivational ones in b). Finally, I will introduce the topic of recalcitrance in section 6, where I also discuss how the views outlined in the other sections fare in relation to the puzzle. The sections follow a somewhat chronological order.⁷

What is an emotion? As Prinz (2004b) points out, there is a significant number of elements one can associate with an emotion, those being: thoughts, bodily changes, action tendencies, modulation of mental processes (such as attention), and conscious feeling (p.3). One way to address the definitional question of emotions, then, is to establish which of these elements is the essential element, which one is identified with the emotion. For instance, some views think the essential element are the thoughts, others the bodily changes and so on. Arguing why one element is more essential than the others is what Prinz calls the 'problem of parts' (pp.3-4), and it garnered a plurality of responses. Another way to answer 'what is an emotion?', however, is to consider all of these elements as necessary constituents. This solution, however, also faces its

⁷ I.e. the sections (except the final one on recalcitrance) are roughly organised in the order in which they were developed and discussed in the literature.

own issues, such as what Prinz calls 'the problem of plenty'. Namely, if all of these parts are essential, then how do they hang together in a coherent whole? Prinz notes the differences between the two problems as follows:

The Problem of Plenty is the counterpoint to the Problem of Parts. The Problem of Parts asks: What components of an emotion episode are really essential to its being an instance of some particular emotion? The tempting answer is that all parts are essential. The Problem of Plenty then asks: If all parts are essential, how do they hang together into a coherent whole? Put differently, the Problem of Parts asks for essential components, and the Problem of Plenty asks for an essential function of emotions in virtue of which they may have several essential components. (p.18)

The problem of plenty is dealt with better by so-called 'hybrid accounts', which argue that essential to an emotion is a particular combination of some elements. In this chapter I will primarily focus on non-hybrid-accounts, but I will mention hybrid accounts that relevantly add to the views I will discuss.

At this juncture it is also worth noting that generally the emotions are seen as divided between basic emotions (fear, anger, joy, sadness), and complex emotions (grief, curiosity, jealousy).⁸ One way of grasping this distinction is evolutionary: our basic emotions have been evolutionarily encoded and inherited, such that we share them with non-human animals. Complex emotions, on the other hand, seem to require a more advanced cognitive system, such that we might think of them as exclusive to humans. I will not rely on this distinction; however, I will indicate the points in the discussion where it could serve as a useful explanation.

2.1 Feeling theories

The most straightforward way of thinking about emotions is to focus on their most phenomenologically salient feature: their felt quality. Thus, feeling theories see emotions as a class of feelings, which we can differentiate from other sensory experiences through their experienced quality. As Scarantino and de Sousa point out, there are two main ways to interpret this type of subjective experience: one is to regard feelings as 'primitives without component parts', the other is to regard them as more

⁸ See Prinz (2004b, pp.86-102) for a discussion of different views of basic and non-basic emotions, leading to interestingly different divisions of emotions into these categories.

complex constructs. When it comes to the former, many philosophers such as Plato, Hume, Locke, have taken feelings to be primitive. However, the latter has received more attention with the rise of the James-Lange theory (Scarantino and de Sousa, 2021).

The James-Lange theory states that the emotions are feelings of bodily modifications ‘triggered by [...] apprehension of certain objects or facts’ (Deonna and Teroni, 2012, p.63). Thus, the emotions are argued to consist in the subject’s internal awareness of bodily responses that result from their interaction with the environment. For instance, being in danger may give rise to bodily responses such as increased heartbeat, sweating, etc., and the awareness of these changes *is* the fear. James bases this result on a thought experiment called ‘the subtraction argument’. According to this argument, if we remove all the bodily sensations from an emotion, we will conclude that nothing remains, i.e. we cannot have an emotion without having the bodily sensations. In what follows I will discuss some of the problems for this theory alongside its advantages, and then move on to some alternatives.

2.2 Advantages and disadvantages of feeling views

A first point in favour of these views is that, as Prinz highlights, there seems to be a link or a correlation between emotions and bodily perturbations. He notes how every culture seems to have bodily expressions for talking about emotional states ‘from broken hearts in our own culture to bad intestines in Tahiti’ (2004a, p.45), which is also corroborated by empirical evidence pointing to correlations between emotional states and changes in the body. These correlations between bodily changes and emotions point to the sufficiency of feeling these changes for an emotion – they can thus arise without the mediation of appraisal judgments. However, Prinz argues that showing sufficiency is not enough, for one should also prove necessity. To that aim he introduces a modified version of the James-Lange theory, which I will discuss later in this section.

A further advantage of this view is, as Deonna and Teroni emphasise, that it is less cognitively demanding and puts ‘the body at the centre’ (p.64), which is an issue for

the views I will discuss in the following section. In brief, by being too cognitively demanding (i.e. understanding emotions in such a way that the subjects would have to have high cognitive abilities) it would exclude infants and animals from the group of subjects who can experience emotions, even though we have a strong intuition that they do experience emotions as well. Furthermore, unlike judgment theories, feeling theories seem better equipped to deal with the problem of recalcitrance, which I will discuss more in detail in following sections.

When it comes to problems for this view, Deonna and Teroni point out that an immediate one is the following: does it not imply the wrong order of explanation? That is, does it make sense to say that we are sad because we cry, rather than say that we cry because we are sad? This would seem to get things the wrong way around, since we give explanatory priority to the emotion over the bodily reactions in our common understanding of emotional experience. For instance, we appeal to the fact that we are sad (so, the emotion), to explain why somebody is crying, or to the fact that they are afraid to explain why their body is shaking. However, Deonna and Teroni point out that this objection is a non-starter, since it gets James' theory wrong: the relation between the two elements is argued to be constitutive, rather than causal. Thus, it is not the case that crying causes sadness, but rather crying *manifests* sadness (p.64).

A further objection casts doubt on one of the assumptions of the view, namely that all emotions are accompanied by bodily changes. Firstly, some emotions do not seem to have any bodily changes, for instance hope, regret, guilt, loneliness, etc. (or at least they are too subtle for us to reasonably notice within the course of an emotion). Secondly, sometimes there may be bodily changes, but they are very weakly felt, or we may be unaware of them – yet we still experience the emotion. Lastly, it seems that people with spinal cord lesions, who are therefore unable to become aware of their bodily changes, still experience emotions normally (Cannon, 1927).

Prinz attempts to deal with some of these objections as follows. First, he notes that James only talks about primary emotions, so more complex emotions such as guilt may have more complex explanations. Second, he also notes that we do not really

know if there are physical correlates of emotions such as guilt. After all, there often are 'pangs of guilt', or maybe facial reactions such as downcast eyes and lowered chin (pp.49-50). That we cannot conclusively point some out does not mean that there are none.

Regarding the patients with spinal cord injury, some empirical studies do point to a diminished emotional experience resulting from the injury. Further, he also notes, drawing from Damasio's work, that most spinal injuries are incomplete and that information about the body can also travel through the blood stream, such that the point about spinal injuries becomes less pertinent to the discussion. Also from Damasio comes the idea of an 'as-if loop', which essentially states that our brains must be able to experience a certain feeling even without the corresponding bodily reaction being present (p.48). Thus, sometimes we may react as if the bodily change was present, even if it is not really there or we are unable to feel it. So, even if other studies show no difference in emotional experience in subjects with spinal cord injuries, that does not imply that their bodies do not in some sense behave in a Jamesian way.

Another serious objection to this view is the fact that, even if we grant that there are bodily reactions associated with every emotion, it is not clear whether it is sufficient to identify emotions. That is, since there seems to be overlap in the bodily reactions of different emotions, then how come we can identify different emotion types? Would it not be the case that, say, anger and indignation, or joy and Schadenfreude, or dread and fear or grief are the same emotion, given how similar they can be in terms of bodily reactions?

To this kind of objection Prinz notes that, indeed, some emotions are very closely related, thus evincing species relations. Emotions such as anger, indignation etc. belong to common families and will thus share some, if not most of the traits. Yet, of course, this does not help answer how it is that we can distinguish between members of the same emotional family. To this he adds that, even though anger and indignation are similar somatically, their eliciting conditions are different. Thus, indignation is caused by injustice, while many instances of anger are not, therefore allowing us to

individuate the emotions. He notes that James and Lange do not mention this idea, but he thinks it would be consistent with their approach (pp.52-53).

Relatedly, a further objection is that a feeling theory such as the James-Lange one cannot explain the fact that emotions have intentional content: the intentional content is not, after all, our body being in such and such a state, but rather a great height, an assailant, a loud noise and so on. So, as I will outline in more detail in the next section, emotions seem to have formal and particular objects (i.e. they attribute some property to some object), which renders the emotions amenable to rational assessment. For instance, we can think of emotions as being appropriate in a given situation, or of emotions being irrational (e.g. being angry when unjustified). How could the James-Lange theory account for this?

Prinz thinks that James-Lange could find a solution by emphasising the eliciting conditions mentioned in the previous paragraph. According to some theories of mental representation, one could say that mental states get their intentional content in virtue of being reliably caused by something (pp.54-55). Assuming that such views of mental representation are correct, then Prinz concludes that there are some causal connections that confer content. He thus introduces his addition to the James-Lange view: core relational themes.

To begin with, he envisions their role as follows:

If emotions are perceptions of bodily states, they are caused by changes in the body. But if those changes in the body are reliably caused by the instantiation of core relational themes, then our perceptions of the body may also represent those themes. In other words, leading theories of how mental representations *entail* that emotions represent core themes, if they are reliably caused by those themes in the right sort of way. (p.55) [sic]

So, what does he understand by 'core relational themes'? The concept relies on what he calls 'elicitation files'. Briefly, these files contain a group of 'body-change elicitors' (including sounds, perceptions, etc.), united in virtue of their association with a particular theme or property, say 'dangerousness'. That is, loud noises, sudden noises, looming objects are all dangerous and thus united in their potency to elicit bodily reactions pertaining to fear. These mental files are not set, so they can be updated depending on our experiences. Getting admitted to these files, however, may depend

on their similarity to the already existing elements. Thus, on this understanding bodily perceptions represent the relevant themes through reliable causation, so we can say that the emotions have formal objects (that is, they assign properties). Furthermore, they can have correctness conditions depending on whether the formal object identified (for example dangerousness) is present or not. For instance, fear can be said to bear a meaning-conferring causal relation to dangerous things, such that being afraid of something that is not dangerous would be inappropriate.

However, one may wonder: what are they attributing these properties *to*? Are emotions about any particular object in any sense? Prinz briefly addresses this issue as follows:

‘A full answer to this question would have to include details about how mental states are bound in thought. The idea would be that a representation of heights gets coactivated with a somatic perception and linked to it in such a way that the former causes the latter to occur, and the latter wanes when the former becomes inactive. There is no reason to provide the details here, because any theory of the emotions will have to explain how dependencies arise between representations of particular objects and somatic states’ (2004a, p.56)

Thus, he does not provide a straightforward explanation of how emotions can be about objects, yet he points out that this is a far more complex matter that any account of the emotions should address carefully.

We have thus seen a stronger version of the feeling view. Further issues with such theories in general will become apparent in the following section, seen as judgmentalism is directly opposed to feeling theories.

3.1 Judgmentalism or evaluative judgment theories

To begin with, judgmentalist (or evaluative, or cognitive) views can be regarded as standing in opposition to feeling views. While the latter argues for a non-cognitive nature of the emotions by emphasising a certain phenomenology, the former emphasises the cognitive nature of the emotions. Thus, on these views an emotion is a judgement about or evaluation of an object. Yet, as Solomon (2003) cautions, the word ‘cognitive’ is here understood in many ways, such that it may even include what is normally called affect. But before I discuss how that may be the case, it would be useful to begin with a survey into what cognitive views entail.

Cognitive views arose as a response to some of the issues with feeling views. For instance, as Nussbaum notes (2004), what distinguishes a mere feeling from a full-blown emotion is that our emotions are *about* something, whereas feelings merely occur in us.⁹ She even argues that the very identity of an emotion depends on having an object – could one be just angry, without being angry at or about something, or could we grieve if there is nothing to grieve over? Furthermore, the type of object discussed here ought to be an intentional object. That is, ‘it figures in the emotion as it is seen or interpreted by the person whose emotion it is’ (p.188). This interpretative aspect of the intentional object is important to note because otherwise we would fail to explain how it is that two people can have different emotional reactions to the same object, e.g. if the object was a dog, then why would some experience fear, while others do not? The difference, then, is how they are internalised or contextualised in the subject’s perspective.

Note that feeling views would struggle to explain why different people have different emotions about the same object. Here a feeling theorist might argue that two individuals simply have different bodily reactions to the same object. However, they would need to do more work to explain (i) what determines the particular bodily reactions to each emotion, since even twins with nearly identical bodies may experience different emotions when facing the same object and (ii) why it is that one individual could have different bodily reactions to similar objects, say my fear reaction to a snake prior to discovering that its species is not poisonous and my calmness when I witness a different snake of the same species – how come my body reacts differently to virtually the same kind of object? Lastly, she points out that our emotions ‘embody (...) beliefs about the object’ (ibidem). That is, there seems to be a strong connectivity between our beliefs and emotions, such that changing one’s beliefs may result in changing emotional state. One example would be being upset after receiving bad news, but then discovering that the news is false and actually the opposite is true – one’s sadness will vanish together with the previous belief.

⁹ As mentioned at the end of the previous section, Prinz argues that it is not so straightforwardly clear that other views can explain how dependencies arise between representations of particular objects and somatic states, so feeling views may not be as disadvantaged as Nussbaum contends.

So, it seems that cognitive elements (whether they be beliefs, judgments, etc.) are necessary for an emotion (and sufficient, depending on the view). The question, then, is what type of cognitive elements. Solomon usefully sums up some of the differing positions: for authors such as J. Murphy and K. Walton, the necessary element is a belief. For J. Neu it is thoughts, whereas for others it comes in the form of evaluative judgments. There are also views that regard the cognitive aspect to have a rather perceptual form, such as R. Roberts' construals or Calhoun's 'seeing as', but I will address those in the section on perceptualism (2003, p.3). Due to space considerations I will only be addressing some of the stronger versions of cognitivism, i.e. that better deal with the objections raised against cognitivism - for instance, Walton's belief view has more difficulties with the objection concerning infants and animals, since it would require subjects to possess high cognitive capacities; hence I will discuss views that can provide better answers to such issues, starting with evaluative judgmentalism.

One initial worry one might have about these views, particularly belief and thought views, is the following: we may have a thought or a belief that p, yet not experience any emotion. I may believe that there is a bear in front of me yet feel no fear. Evaluative judgment views attempt to bypass this issue by noting that the cognitive content we are after (for them, the judgment), cannot be of a merely factual form. Instead, it needs to evaluate the object - it needs to be of the form 'the bear is dangerous', where the word 'dangerous' is the formal object (or evaluative property) attributed to the object. The fact that there is a bear in front of me, if unqualified in any way, cannot give rise to any emotion in me. Yet judging the bear to be dangerous seems to be a more promising stance.

Thus, the judgments involved are not neutral apprehensions of the object, but they apprehend it in an evaluative way - so, they are intelligible emotional responses. Some evaluative judgment views do not stop here: indeed, even an evaluation in terms of formal objects can occur in the absence of emotional responses. A strategy defended at this point is to argue that emotions are concerned with value, specifically values that the subject has. A good example of this kind of view is Nussbaum's eudaimonistic view (2004, pp.189-190). According to her, such evaluations affect us emotionally if

they are connected with values that are relevant to our flourishing as human beings. Her main example is the experience of grief over the loss of a loved one: we do not grieve the death of a completely unknown person, unlike the loss of our loved ones, because we value the latter in a way that we do not the former. Thus, I may have the judgment that 'Their death is a loss' (assuming 'loss' is the evaluative property at play), yet not feel grief. But I will definitely experience grief if the person is relevant to my flourishing, such as a loved one (so, the judgment will still be 'Their death is a loss', yet in this case tied to the subject's concerns).

A few clarifications are in order. In particular, we need to clarify what kind of mental state are judgments, for one may think it absurd as a requirement that we always have a conscious judgment in a propositional form. Solomon carefully sums up the following traits as pertaining to judgment as he understands it:

'They are episodic but possibly long-term as well. They must span the bridge between conscious and non-conscious awareness. They must accept as their 'objects' both propositions and perceptions. They must be appropriate both in the presence of their objects and in their absence. They must involve appraisals and evaluations without necessarily involving (or excluding) reflective appraisals and evaluations. They must stimulate thoughts and encourage beliefs (as well as being founded on beliefs) without themselves being nothing more than a thought or a belief. And (of considerable importance to me), they must artfully bridge the categories of the voluntary and the involuntary.' (2003, pp.10-12)

I will not examine each of these elements, but I will address two extremely important ones. On the one hand, both Solomon and Nussbaum emphasise the dynamic aspect of these judgments (i.e. that they are not merely passive, short-lived mental occurrences). Nussbaum in particular borrows the Stoic understanding of judgment as 'assenting to how the world is'. That is, the judgment includes recognising its content, and rejecting or assenting to it respectively. It does not suffice to merely entertain the thought - one must assent to it, take it as true (2004, p.192). If I am walking through a hologram museum and I see a bear in front of me, I do not necessarily assent to the thought that there is a bear in front of me - I take in the elements and reject them, hence I am not afraid. Of course, assenting to them (so taking it as true that there actually is a bear in front of oneself) would typically lead to fear. This may lead to several complications from the point of view of recalcitrance, but I will not address these at this point. The main aspect to note here is that one is constantly evaluating the situation, such that a judgment is not merely an episodic

occurrence: we do not judge 'X is p' for one second and then stop making that judgment. So long as we are still inhabiting the situation, we are still actively assenting to a certain state of facts, e.g. to the presence or absence of the bear.

On the other hand, judgments need not be conscious. In fact, many of them may be judgments we are not aware of, given the high-paced reactivity of some emotional episodes. At this point it is worth introducing Solomon's discussion of the word 'cognitive'. First, he discusses the apparent conflict between the two sides of the debate: feeling and cognition. He notes that the feeling side uses the slogan 'Emotion before cognition', to signify that we first experience a particular emotion and only then do we form certain judgments. However, he questions the very basis of the feeling/cognition dichotomy: to have a response to something, particularly an emotional response, there first needs to be some type of recognition (of whatever it is we are responding to). So, even if we may start from this minimal sense, recognition is a type of cognition and thus we should not distinguish so sharply between these elements. Furthermore, he adds that there are primitive pre-conceptual forms of cognition, so when we endorse cognitive views, we are not thereby confining ourselves to articulate cognition (p.2).

A final important thing to note from his discussion of 'cognition' is the fact that it need not exclude feelings. Indeed, one of the major flaws of the cognitive view has been the fact that it seems to completely leave out feelings, or the felt quality of our emotions. Solomon thinks that, while in other works he may have been guilty of this himself, he now understands cognition in a way that can incorporate feeling. Specifically, he notes that a significant amount of what is 'unhelpfully called affect' can be identified with the body, in particular with 'the judgment of the body', a concept he draws from Downing's work (2001). By that he means some of our 'body micropractices', explaining that a lot of our cognition is of a pre-linguistic nature (and not a pre-cognitive one, as it often is mislabelled). While Solomon does not offer more detail regarding how 'body micropractices' are to be understood, he does offer two arguments to the effect that they can be meaningfully referred to as 'judgments'.

The first one is a point he has already established: namely that judgments are not necessarily articulate or conscious, so there would be no problem in calling such bodily phenomena 'judgments'. The second one indicates how we are creatures who admit of both knowledge that and knowledge how, so our bodily judgments may well be of the latter type. That is, some of our emotional reactions can be seen from the point of view of habits and practices we perform, rather than ways of evaluating the world. He then cashes these out in terms of feelings of comfort and discomfort: for instance, shame is at least in part a feeling of discomfort with other people, of being rejected. That felt discomfort is, then, a bodily judgment to the effect that we in some sense do not belong with our group. So, this is an example of how judgmentalist views could incorporate feelings into their view, a challenge that will be addressed more in detail in the following section.

3.2 Advantages and disadvantages of judgmentalist or evaluative accounts

Deonna and Teroni (2012) carefully list some of the major benefits of a judgmentalist account. Firstly, it constitutes an effective way to organise the class of emotions, since it allows us to individuate emotion types in virtue of their evaluative properties. Thus, all instances of fear can be categorised as fear in virtue of them characterising their object as dangerous, anger in virtue of characterising them as unjust, etc. This goes hand in hand with another advantage, namely that it helps us specify the correctness conditions for an emotion. Emotions may be appropriate or not in a certain circumstance, yet other views (say, a feeling view) would be unable to tell us why. Judgmentalism, on the other hand, presents emotions as evaluations, which means that there is a straightforward way in which they may misevaluate their object. Thus, if misfortune strikes a joyful response in somebody, their emotion is inappropriate because it evaluates the situation as joyful, when in fact it is not a warranted evaluation (p.53).

Furthermore, judgmentalist views satisfy the intuition according to which emotions have a mind-to-world direction of fit; that is, emotions aim at matching the world, rather than making the world fit the mind. Lastly, these views can account for the links

between emotions and motivation, since we can track a clearer connection between a judgment of the existence of a property (e.g. dangerousness) and a motivational reaction to it (running away from the object) (ibidem). When cashing out this connection, however, Deonna and Teroni discuss these two options:

A first option is to endorse motivational internalism, according to which evaluative judgments are essentially motivating. This option seems problematic, since it is not straightforwardly true: one may evaluate something a certain way without being motivated to react to it in a particular way. Additionally, this kind of approach would make it difficult to account for the emotions that are not intimately connected to actions, for instance those that are directed at the past. An example would be regretting something one has said many years ago, even though the issue has long been sorted – the kind of shame that suddenly overcomes one when thinking about such long-gone instances. The second option is thus more plausible: the emotions can be regarded as reasons for certain desires without standing in a constitutive relation with them. Those emotions that are not intimately linked with action ‘will for their part not engender any desire’ (p.54).

Deonna and Teroni also discuss some serious objections to the view. Their discussion being a rather lengthy one, I will restrict myself to only the main issues and potential solutions to those. The main issue with judgmentalist views, then, is that evaluative judgments are neither necessary nor sufficient for an emotion. When it comes to the necessity claim, there are two ways to argue against it.

On the one hand, not having that judgment seems to be a common occurrence, as observed in cases of recalcitrance: my conscious evaluative judgment is that the spider is not dangerous, so how can it be that I both think it dangerous and not dangerous (ibidem)? One way to circumvent this issue is to conceive of the emotions as less committed types of attitudes towards the relevant proposition, for instance in the guise of construals or seeings as, which I will address in section 4. I will return to the issue of recalcitrance in section 6.

On the other hand, it also seems to be too demanding with regards to the cognitive capacities expected of the subject, since they must master the concepts that figure in

the evaluations they hold. Animals and infants are thus a category of subjects who would fail to satisfy this criterion, yet we would still want to regard them as having emotions (p.55). However, it must be noted that judgmentalists have responded to both of these objections. Solomon, for instance, explicitly holds that the judgments need not be articulate or conscious, often manifesting itself in primitive, pre-conceptual formats. Thus, he can accept that infants and animals may not have the same wealth of emotions as human adults, yet that does not mean they do not experience the more basic emotions (2003, pp.2, 16).

When it comes to the sufficiency claim, we have briefly touched on it in the previous section: one can have an evaluative judgment without experiencing an emotion. Hence, it seems that something is missing, and that something seems to be the phenomenology. As we have already seen, not every judgmentalist would accept this as an issue: for Solomon, the feeling is already in the judgment. Furthermore, we can think of feelings outliving the emotion and of emotions outliving the feeling – thus, it is questionable whether they are an essential feature on their own (p.13).

At any rate, the issues surrounding the sufficiency claim have led to new approaches within judgmentalism, so called add-on strategies. Thus, theorists would include a new essential component to the emotion, namely a certain phenomenology. This is a good example of a hybrid theory, where more than one element is considered necessary for an emotion. An example of such a judgment plus feeling theory would be Aristotle's, for whom emotions are a certain type of evaluative judgment accompanied by pleasure or pain. For instance, in the *Rhetoric* he defines anger as: 'a desire accompanied by pain, for a conspicuous revenge for a conspicuous slight at the hands of men who have no call to slight oneself or one's friends' (1378a31-32). As Deonna and Teroni highlight, this solution is not free of difficulties. Particularly, it raises questions about the roles it assigns to the phenomenology and the evaluative element, which harkens back to Prinz's 'problem of plenty' – what relation do these parts stand in (2012, p.57)? I will not discuss the possible solutions due to space considerations.

Finally, Deonna and Teroni mention a different type of add-on strategy, which embodies the reverse of the strategy in the previous paragraph. Specifically, instead of adding feeling to the judgment component, some theorists such as Barrett (2006) add the judgment to the feeling element. That is, they argue something briefly as follows: first, we experience certain bodily feelings, that we then interpret via judgment in light of the circumstances, such that the interpretative act is what turns our initial feelings into full-blown emotions. These views (which are a hybrid version of judgmentalism) are more common within psychology, and I will not discuss them due to space constraints.

4. Perceptualism

Perceptualism emerged as an alternative to judgmentalism and feeling theories, since it is considered a third way between them that 'promises the best of both worlds without the perils of either' (Salmela, 2011, p.1). In other words, Salmela notes that the attractiveness of this view is based on the fact that perception is a (*prima facie*) plausible notion 'for combining the most conspicuous dimensions of emotion, *intentionality* and *phenomenality*, in a way that is capable of escaping standard objections to both traditional cognitivism and noncognitivism' (p.1, original italics). So, in an attempt to characterise the emotions in a way that does justice to both their intentionality and phenomenality, a promising alternative emerged by borrowing from the philosophy of perception.

Before outlining how exactly one is to understand the recourse to perception, Deonna and Teroni highlight some initial merits of such views. In the first place, perceptual theories maintain the direction of fit generally thought to characterise the emotions: mind-to-world, i.e. aimed at describing the world, without assimilating the emotions to judgments, which we have seen to have a series of drawbacks. These views thus satisfy the intuition that the emotions play the important intentional role of 'revealing a world of values' (the values they reveal being normative values, i.e. what has so far been called 'formal properties', for example 'dangerousness' for fear). Related to this advantage is also the fact that perceptual views would thus allow us to individuate

emotions based on the values they reveal, which was a benefit of judgmentalist views that feeling theories struggled to account for. Further, perceptualism also respects the intuition about the phenomenology of emotions according to which ‘emotions are a form of affective perception through which a world of evaluative properties is presented to us in an expertly salient way’ (p.67).

Lastly, this view is also promising when it comes to resolving the challenge of recalcitrant emotions, for they could be likened to perceptual illusions. Just like I may know that the stick is straight, but perceive it as bent, I may also be afraid of a spider while thinking it is not dangerous (p.67). There will be a more extensive discussion of a perceptualist approach to recalcitrance in section 6, as well as in chapter 2, section 2.2.

However, how should one understand this recourse to perception? Is the claim that emotion *is* a type of perception, or rather merely *like* perception? There has been support for both views, the stronger and weaker types of perceptualism respectively. I will first discuss the strong view or, as Salmela calls it, ‘emotion as a perceptual system’ (2011, p.4). Then, after discussing reasons against such a strong view, I will move on to the weaker one, which argues that emotions work analogically to perception.

a) Strong perceptualism

To begin with, the stronger claim argues that ‘emotions are, in essence, perceptual experiences of evaluative properties’ (Tappolet, 2016, p.15). Or, as Salmela puts it, this view suggests that ‘emotions constitute a distinct perceptual system that resembles traditional sense modalities in all relevant aspects’ (2011, p.3). An example of such a view, and the one Salmela discusses in detail, is Prinz’ view introduced in section 2.2. I will not discuss his view again, yet I will note that his view can be considered a strong perceptualist one in virtue of the fact that to him, the emotions are essentially perceptions of somatic changes, with the addition that they also highlight matters of concern for the subject in virtue of their function of being reliably caused by our respective elicitation factors (for Salmela’s full discussion of Prinz, see pp.4-11).

Deonna and Teroni offer a list of compelling reasons against the strong view, of which I will mention the most noteworthy ones. An initial worry may be the fact that actual perception is associated with sensory modalities and organs, whereas emotions are not associated with any such bodily systems. There is, in other words, no emotional organ like the eye, nose, etc. (2012, p.68). However, Tappolet notes that this argument is not as problematic as it may seem at first. She begins by emphasising that whether specific organs are necessary for something to count as perception will depend on the view of perception one espouses. Then, should one still be inclined to consider sense organs a crucial aspect of perception, she notes that this would then exclude proprioception (the perception of our bodily movements and special orientation) from the ranks of perception. Further, even if proprioception is aided by joint receptors, these are not the same as sense organs. Lastly, she also argues that multimodal perception would struggle to be accounted for by such a rigid view of perception. As an example, she notes that speech perception may involve both vision and audition without having a dedicated sense organ. Thus, she concludes that 'that emotions do not depend on organs is only a reason to think that they are not sensory experiences; it is not a reason to think that they are not perceptual experiences' (2016, p.28).

A further objection is that the phenomenology of emotions, despite being *prima facie* closer to perception than to judgment, turns out to not be as easy to assimilate to the phenomenology of perception. First, emotions are valenced experiences, unlike perceptions, which are not essentially so (Deonna & Teroni, 2012, p.68). That is, emotions have either a positive quality (joy, elation, etc.), or a negative one (anger, sadness, etc.). Perceptions, on the other hand, are neutral – it is neither positive nor negative to see a vase qua perception of a vase. Whether one has personal connections to that vase such that the perception takes more of a positive valence, that is outside of the raw perception we are interested in. Anger is inherently negative, whereas perceptions such as seeing a table are not inherently valenced, though they may gain a valence depending on one's other mental states.

Connectedly, there is another concern related to the phenomenology, namely that perceptual experiences are transparent, unlike the emotions. What transparency

means in this context is that the properties of the perceptual experience seem to be reducible to the properties of the object perceived. In other words, one cannot attend to a perceptual experience one is having without attending to the objects and properties perceived. For instance, if one attempts to describe one's perceptual experience of a vase, the description will be limited to the properties of the vase itself (it is blue, it is small, etc.), rather than properties specific to having a visual experience. When it comes to the emotions, on the other hand, describing the experience of fear will involve describing properties of one's emotional state, not necessarily those of the object. When I describe my fear, I describe what it feels like to me, I do not describe the spider in front of me (assuming, of course, that my aim is to convey what kind of state I am in, rather than leading the hearer to experience a similar state for themselves) (p.68).

Another issue with strong perceptualism is the fact that, as Deonna and Teroni highlight, perception requires its object to be present, whereas emotions do not. That is, we can have emotions directed at the past (say, nostalgia for how things used to be), at the future (being afraid for the future of our children), or even at fictional objects (being sad for Anna Karenina). On the other hand, we cannot perceive something that is absent (p.70). This objection will also turn, as Tappolet argued, on what conception of perception we choose to work with. As she claims, the mere fact that there are differences between (other) perceptions and emotions is not enough to rule them out from being perceptions (2016, p.28).

Now we can turn to what Deonna and Teroni deem 'the most serious objection' to strong perceptualism. Specifically, even though the emotions can access the evaluative properties they reveal, they are not independent ways of accessing the objects that exemplify those properties. For instance, while one's fear correctly identifies the dangerousness of the bear, it can only do so on the basis of first perceiving the bear via vision or audition. Alternatively, the offensiveness of a remark can only be identified after one remembers, hears, considers etc. the remark itself. In other words, one needs to rely on a cognitive base in order to access the evaluative properties of emotions, whereas perceptions simply identify their objects without mediation. Thus,

such cognitive bases cause and, in some sense explain the occurrence of emotions, whereas there 'is no such comparable distinction between two psychological levels exemplifying causal and epistemic relations within the field of perception proper' (p.69).

Regarding the issue of cognitive bases, Tappolet offers a subtle response: sure, the emotions are not as 'open' to the world as perception, since they require a mediator to access it (cognitive bases). However, she urges us to look at the times when 'things go well', by which I take her to mean the situations in which emotions do not misfire (2016, p.30). Unlike perceptions, which misfire mainly due to their hardware (i.e. biological impediments), emotions importantly also misfire because of the cognitive bases (say, one falsely believes immigrants are evil and is therefore afraid of an innocent-looking foreigner). So, when things do 'go well', and our cognitive bases function properly, then why should the fact that they are mediated count as a reason against the openness of emotions?

Likewise, even if it seems odd that the emotions would afford us openness to things that are not present (past or future objects), unless one has an overly strict causal constraint on perception, then what is the problem with saying that the emotions afford us openness to non-present entities? In other words, her move is to shift the burden of proof: the anti-perceptualist must now explain why a different level of openness, even if mediated and therefore more susceptible to misfiring, is therefore to be excluded from the category of 'perception'. Yet, to appease the anti-perceptualist, Tappolet suggests an amendment: emotions could be seen as 'quasi-perceptions', rather than full perceptions - even though she notes this is more of a terminological, rather than substantial issue (ibidem). We will thus move on to weak perceptualism.

b) Weak perceptualism

As already mentioned, weak perceptualism proceeds by analogy: by highlighting some similarities between the emotions and perception, we develop a clearer picture of what the emotions are. In this section I will briefly discuss some analogies and

disanalogies, and then introduce Roberts' construals account as an example of what a well-developed weak perceptualist view looks like.

First, I already noted some important similarities between perceptions and emotions in the previous section. As such, I will now briefly enumerate several other similarities, as listed by Tappolet (2016). First, they are both conscious states characterised by phenomenological properties. Second, they are automatic, in the sense that they are not directly subjected to our will (so, in other words, they are passive states). Third, they have correctness conditions: my perception of a blue vase is correct just in case the object I am perceiving is both blue and a vase, and my emotion is appropriate depending on its circumstances, such as grief at a funeral (pp.19-20).

Fourth, they both seem to be informationally encapsulated, i.e. other mental states have a limited impact on them (not in the sense of influence, as cognitive bases of course can be said to influence our emotions), but rather when our perception or our emotion is in competition with other mental states, it seems that both emotions and perceptions 'win' (or at least do not lose). That is, even though I know that the stick is straight, my perception remains unchanged, while even though I know that the spider is harmless, I will remain afraid (p.20).

There are, of course, other analogies, but I think these are the most noteworthy ones. Then, Tappolet also mentions a few disanalogies, some of them already discussed: the absence of analogous organs, the more complex phenomenology of emotions, the fact that they rely on cognitive bases, as well as other psychological factors such as moods, desires, etc., the fact that they are valenced, and the fact that perceptions are transparent. Furthermore, emotions are closely linked with motivation and action, unlike perception.

Lastly, and importantly, there are disanalogies concerning the rationality of emotions. First, experiencing (sensory) illusions (such as the Müller-Lyer illusion) is not seen as

irrational, while experiencing recalcitrant emotions is seen as irrational.¹⁰ I will outline this issue in more detail in section 6, the section dedicated to recalcitrance. Second, we can ask why-questions about our emotions, but not our perceptions. Lastly, emotions grasp normative things (evaluative properties), whereas perception only grasps at properties such as shapes or colours (p.27).

I will now briefly discuss one view that falls under this category: Roberts' (1988) view of emotions as construals. According to Roberts, emotions are concern-based construals. First, by construals he understands a 'mental event or state in which one term is grasped in terms of something else' (p.190).¹¹ To better grasp this claim, he talks about construing in relation to Wittgenstein's 'seeing as': 'construing seems to involve dwelling on or attending to or holding onto some aspect' (p.187). For instance, he discusses the duck-rabbit picture and how one is able to construe it both as a duck and as a rabbit; that is, one is able to perceive the picture in two different ways, depending on whether one focusses on the 'duckiness' or the 'rabbithness'. Likewise, the emotions are ways of construing objects in certain way, say as dangerous or offensive.

However, what makes it such that one construes objects one way over another, say dangerous over friendly? According to Roberts, the crucial aspect seems to be that concerns are involved. By 'concerns' he means desires and aversions, alongside the attachments and interests that many of our desires and aversions derive from (p.202). So, a concern about my safety will make it such that I will construe the bear as dangerous, rather than friendly. Concerns can be either biological (survival), or learned (wanting to study well), general or specific (wanting to do well in business or

¹⁰ Tappolet later provides an explanation for why this might be: briefly, the plasticity of our emotional dispositions makes it such that we ought to work on them. As such, it makes sense to apply rational criteria to our emotions, since it would lead us to identify whether there is something wrong with our emotional system (2016, p.38). However, Döring offers a compelling argument against this explanation, which is that even if one has, say, a fear of heights, as well as a desire to go hiking, this person is under no rational requirement to work on their fear of heights – they may simply find a new hobby (2014, p.127). Döring then provides an alternative explanation, which I discuss in Chapter 2, section 2.2.

¹¹ He clarifies that the 'in terms of' relation can have as items perception, thoughts, images, concepts. As Tappolet notes in fn.37, Roberts is mistakenly taken to be a quasi-judgmentalist. His view counts as perceptualist, however, since 'construals can be non-propositional and perceptual in a broad sense' (Tappolet, 2016).

wanting to do well in business x), ultimate or derivative (attachment for a family member can be seen as ultimate, since it would provide reasons for other concerns, such as taking care of them), and dispositional or occurrent (one can be disposed to be concerned about keeping one's home fire-free, and one can have an occurrent concern about it by keeping it in mind at a particular time) (ibidem).

Lastly, concerns must be serious. That is, they must be verisimilar in the sense that they have the appearance of truth for the construer, such that they do not necessarily affirm that construal, yet still experience it (p.201). Say, if I believe that a child does not know what they are saying when they are simply copying some offensive words, I will not construe their act as offensive because it does not appear as an offence to me.

We have thus seen one example of a weak perceptualist position. I will only highlight one issue with this view, since it bears on our discussion at large: the way it deals with recalcitrance. For Roberts, the difference between our reaction to perceptual illusions (not regarding them as irrational) and our reaction to recalcitrant emotions (seeing them as irrational and attempting to distance oneself from them) is due to the fact that emotions essentially involve concerns. Since our fear is permeated with deep-rooted personal concerns, attempting to distance ourselves from our emotion is like rejecting one part of ourselves, whereas there is no such inner conflict when it comes to perceptions. However, as Brady notes, this solution is not as helpful as it may seem, for, surely, we also have a concern that our visual perception represents the world accurately. The appeal to concerns to explain the gap between recalcitrant emotions and perceptual illusions thus seems to fail (Brady, 2007, pp.276-277).

5. Attitudinal and Motivational views

The other significant strand in the philosophy of emotions is that of attitudinal and motivational views. As Scarantino (2014) highlights, emotions include an appraisal, a feeling, and a motivation, yet the majority of philosophers begin their inquiry from either of the first two. Consequently, judgmentalist and perceptualist views have dominated the discussion. However, Scarantino goes on to say that neither of these views can explain how emotions motivate us to act: there is no straightforward

connection between perceiving the world to be a certain way, or believing it to be that way, and taking a specific course of action. We therefore should consider beginning our inquiry from the motivational component as well (p.157).

According to Scarantino and de Sousa (2021), there are two main types of views within the motivational tradition: on the one hand, there are phenomenological views, such as Deonna and Teroni's, which equate the emotions to feelings of action readiness; on the other hand, there are non-phenomenological views, such as Scarantino's, which equate the emotions to 'causes of states of action readiness which may or may not be felt'. What they both agree on, then, is that 'the fundamental aspect of an emotion is the way it motivates the emoter to act'. In this section I will first provide a brief overview of Deonna and Teroni's account, followed by a summary of Scarantino's.

a) The Attitudinal View

Deonna and Teroni (2012) begin by clearly distinguishing between attitudes and contents. Attitudes such as belief, desire, supposition and conjecture can all be about the same state of affairs. That is, one may believe that *p*, while another person merely supposes that *p* - since believing and supposing are different attitudes, the formal objects are different as well. The formal object of belief is truth (so, 'to believe is to take a proposition to be true'), and possibility is the formal object of supposition (supposing, then, is 'to take a state of affairs to be possible'). Thus, even though the content of these mental states is identical (*p*), they are nevertheless different mental states in virtue of their differing attitudes. Further, the conjunction of attitudes and content is what sets the correctness conditions of an attitude: since the formal object of belief is truth, the correctness conditions of my belief that *p* is for *p* to be true. If I suppose that *p*, the correctness conditions change to *p* just being possible (pp.76-77).

Emotions are also a type of attitude for Deonna and Teroni. If two people have different emotional reactions about one and the same object (say, a bear), the difference between their emotions is not to be located at the level of content, for that would imply their emotions are not about the same thing. One person would emote towards, say, a specific instance of dangerousness (if they are afraid), while another would emote towards an instance of funniness (if they are amused). Since we want to

say that the two emotions are about the same thing, it would make more sense to say that they are different attitudes towards the same content. Note, however, that this would imply that the formal object (dangerousness, etc.) is not in the content of the emotion. So, if evaluative properties are not in the content of the emotion, how are we to understand emotions as evaluative attitudes (p.77)? This will become clearer in what follows.

First, they clarify that there is not one main emotional attitude, emoting, which would be the attitude of taking something as (dis)valuable. They contend that there is no reason to think that such an attitude would have a psychological reality over and above that of its determinate instances. So, they conclude that, for each type of emotion, there corresponds a distinct type of evaluative attitude. They here also offer a quick taxonomy of the different types of attitudes: propositional or non-propositional (such as episodic memory, sensory imaginings), occurrent or dispositional, as well as with or without a salient phenomenology (p.78).

In order to obtain a definition of what they take emotions to be, they look at a series of constraints for any such definition: emotions are often not directed at propositions, they are episodes, they have a salient experiential dimension, their phenomenology is best captured in terms of bodily feelings, and it is in virtue of their phenomenology that emotions relate to evaluative properties. Given these considerations, they conclude that 'the emotionally relevant bodily changes are experienced as distinct stances we adopt towards specific objects' (p.79). So, they argue that we should conceive of the emotions 'as distinctive types of bodily awareness, where the subject experiences her body holistically as taking an attitude towards a certain object' (ibidem). Thus, for them emotions are intimately connected with types of action readiness or felt action readiness.

Drawing on Frijda's work (1986), Deonna and Teroni emphasise how action readiness imbues the world with significance, i.e. that states of a (felt) readiness contribute essentially to the world being presented to the subject as significant in various ways (p.80). However, this does not mean that emotions are attitudes we take towards what we feel happens in our body, but rather that what we feel happening in our body is

itself an emotional attitude. To further clarify, by 'action readiness' they mean a broad range of actions we are ready to undertake, not just fight, flight or freezing. Even more minor tendencies such as focusing on an object rather than on another, smiling, blinking nervously, are all types of action readiness (ibidem).

Given this understanding, we can now ask the question again: in what sense do these states of felt action readiness make the emotions evaluative? In other words, how do these distinct felt bodily stances help account for the fact that emotions are only correct if their object exemplifies a specific evaluative property? To answer this question, Deonna and Teroni point out the following: our action readiness in any given circumstance (say, flight in the face of a bear) is correct if and only if the object does exhibit the evaluative property we associate with that emotion (p.81). That is, feeling oneself ready to run away from a bear, an action that would diminish the bear's impact on oneself, is correct if and only if the bear is dangerous. Thus, we can say that fear is experiencing the object as dangerous because of this intricate connection between the action tendency and the evaluative property necessary for making that particular action tendency correct or incorrect.

Before moving on to virtues of the theory, Deonna and Teroni make two observations: one, their approach treats bodily feelings as constitutive of the emotion, not as causes (the difference between their approach and James' being that their account illuminates how emotions are directed towards the world); two, emotions are stances towards objects, but not necessarily bodily attitudes. For instance, regret and pride are instances of felt attitudes that are not embodied. However, they think that it might still be possible to account for such emotions in terms of felt, yet not embodied attitudes. They suggest that their view has a more flexible architecture than the Jamesian one, which could facilitate such an endeavour. Furthermore, throughout their book they have questioned the viability of non-embodied accounts of emotions, which should foster optimism for their project (p.81).

As Scarantino and de Sousa (2021) highlight, this is the biggest challenge for both motivational and attitudinal theories: accounting for the states of action readiness distinctive of different emotions. Many emotions do not motivate actions at all, such

as grief, depression, or even backward-looking emotions such as regret. Furthermore, there are emotions such as joy that have an open range of behavioural options, such that it would be difficult to pinpoint a particular action tendency that should be associated with it (and therefore fix its correctness conditions). Lastly, there is also overlap between the actions associated with different emotions, such that it becomes difficult to argue that one action pertains to one emotion over the other. So, Deonna and Teroni's contention that their view could find a way to account for such emotions seems to face a slew of difficulties.

However, they also note some of the virtues of their view. Unlike most feeling views, their account can individuate emotions, since the felt bodily attitude towards an object is only correct iff the objects exemplify a given evaluative property. This view can thus explain the variety of emotion types. Furthermore, it helps account for how subjects are motivated to act in specific ways and form specific desires. Specifically, they address how certain desires associated with certain emotions (such as the desire to climb up a tree when you see a bear) are best understood as caused by the emotion, rather than constituents of them. An additional advantage associated with this claim is that their view 'illuminates many of the relations between emotions and evaluative judgments', for example why an evaluative judgment is thought to imply that we are motivated to act in accordance with that judgment (p.83).

Relatedly, they state that 'the hallmark of a sincere evaluative judgment is being in the corresponding motivational state' (p.84). Thus, their view is better than other evaluative views that must commit to some type of motivational internalism. Another virtue of their view is that emotions are not about evaluative properties, which they consider a negative implication of perceptualism. Rather, emotions have evaluative conditions of correctness in virtue of being the attitudes they are.¹²

A further virtue of Deonna and Teroni's view is its ability to explain animal and infant emotions: they are able to feel their action tendencies, and so to have the particular emotional attitudes. What 'they do not have', add Deonna and Teroni, 'is the capacity

¹² See Rossi and Tappolet (2019, pp.550-551) for an argument against Deonna and Teroni's claim that the formal objects of emotions do not feature in the content of their tokens.

to build on their emotional responses so as to come to understand that only some situations make them appropriate' (p.85). That is, they cannot yet make evaluative judgments, yet they can experience the emotion.

Lastly, and crucially for our purposes, they add that this explanation also provides us with the resources to address cases of irrational emotions (such as recalcitrance). It is, according to them, not difficult to see how one could come to adopt a given emotional attitude towards a certain object while aware that the situation in question does not make it appropriate. Not only that, but when compared with perceptualism, they note that their view offers a better account of emotions such as phobias or other 'emotional errors'. According to perceptualism, phobias, like perceptual illusions, are merely mistaken but are not irrational. On the attitudinal view, however, they 'might well be fully irrational in the sense of unjustified' (p.85). In other words, their account preserves the intuition that recalcitrant emotions are in some sense irrational (they fail their correctness conditions), unlike perceptualism, which arguably replaces the irrationality with faultiness akin to that of perceptual illusions.

b) The Motivational View

As mentioned at the beginning of this section, Scarantino (2014) offers a non-phenomenological motivational account of the emotions. His starting point is the observation that three key features of emotional motivation are unaccounted for by competing accounts: impulsivity, flexibility, and bodily underpinnings. He therefore develops the 'Motivational Theory of Emotions' (henceforth MTE), which he shows to be able to account for these features. Further, he argues, MTE is also able to deal with the problem of intentionality and the problem of differentiation (p.156).

First, what are these three key features? To better illustrate the relevance of these features, he solely focuses on 'the most prototypical of emotional actions', i.e. impulsive emotional actions. These are intentional acts done in the heat of the moment, such as slapping someone out of anger. The first key feature is impulsivity, which has the following characteristics: it has a sense of urge, which encompasses both an expectation for gain after completion and a haste to fulfil it; additionally, it leads to

us using only one part of the available cues that might indicate the adequacy of action, property which he calls short-sightedness. Short-sightedness is due to two potential sources, either reduced information gathering (since we are experiencing a sense of urgency, we will not have time or the drive to consult all the available sources of information), or information processing biases (pp.157-159).

The second key feature is flexibility, which stands for the fact that the same emotion can lead to a variety of different actions, depending on the circumstances. That is, we may run or hide in fear, or call the police, but we will not stop to take a nap or read a book, which shows that this flexibility is also constrained. Importantly, Scarantino also notes within this context how an account will have to also explain why some emotions do not lead to actions, while some even lead to actively refraining from action (such as grief or depression). The third key feature is that of bodily underpinnings, which means that when we act emotionally, we also have certain bodily changes we do not choose voluntarily, such as facial, vocal, and postural reactions (pp.159-160).

He then proceeds to show that neither cognitivism nor perceptualism can make sense of emotional actions, particularly of impulsive actions. However, due to space considerations I will not explore his arguments here, but I will instead discuss how his view accounts for this type of motivation.

His view relies on Frijda's (1986) conception of emotions as action tendencies, understood as states of readiness to execute a given kind of action, specifically the kind of action defined by the end result aimed at (in other words, a relational goal). In Frijda's conception, the sorts of action tendencies that emotions should be identified with have control precedence. That is, that particular action tendency 'seizes control of the emoter with respect to their mental or physical actions' (p.170). Those action tendencies acquire control precedence in virtue of two functional components. The first one, *precedence*, states that the action tendency takes precedence over other actions and states of action readiness and is also manifested by the fact that they interrupt other processes. The second one, *preparation*, states that we are not simply ready for action, but preparing for it. Furthermore, he points out that these processes tend to come in degrees, such that we can obtain higher or lower degrees of precedence and

preparation. Note this means that there may be some mental states that may fail to fully qualify as emotions or as non-emotions (ibidem).

Before accounting for impulsive emotional actions, he addresses the challenge of emotions that reduce willingness to act, such as grief and depression, as well as the challenge of energising emotions, such as existential joy, which do not have any particular relational goal. His solution to these challenges (which, as mentioned at the beginning of section 5, are the biggest challenge for motivational views) is to argue that the relational goals of joy and sadness are significantly less specific than those of other emotions such as fear. For instance, the relational goal of joy is to relate as such, to be creative, to play. As he puts it, 'the joyful person is ready to engage in an open range of actions, and actively prepares for this open engagement with the world with a generalized state of arousal' (p.171). The goal of sadness, on the other hand, is to not relate as such, or to disengage from the world in an undifferentiated fashion, in the sense that there is not much they would like to do. So, he concludes that there are both focused and unfocused action and inaction tendencies, and they all instantiate control precedence insofar as they are emotions (ibidem).

He then provides a thorough explanation of emotional impulsive acts. First, he states that emotional acts result from a two-level structure of control. The first level is constituted by the action and the inaction tendencies with control precedence, which are special purpose motivational structures. The second level is constituted by a set of general-purpose capacities, such as rational control, which determines if and how the prioritised (in)action tendency is manifested.¹³ There are two types of rational control that occur at the second level. On the one hand, there is compatibility control, which monitors whether the emotional relational goals and sub-goals are compatible with the emoter's other goals and value systems. For instance, reacting aggressively towards one's boss would conflict with one's goal of career security. On the other hand, there is executive control, which secures that the emotion's relational goal is translated into a set of sub-goals that is instrumentally adequate. For instance, once

¹³ The basis for this two-level differentiation is Gallistel's (1980) empirical treatment of action, which points out that such structures of control are the primary means to achieve flexibility in the animal kingdom.

one has decided to slap one's partner in anger, executive control ensures that the motor system and specific body motions involved in that action can happen successfully (p.172).

So, how does this two-level control structure explain impulsive actions? According to Scarantino, 'impulsive emotional actions come about whenever the interaction between rational control and a prioritized action tendency quickly leads to an intention-in-action, namely an intention to act now so as to achieve the relational goal of the action tendency' (p.173). In his words:

'The impulsivity and bodily underpinnings of prototypical emotional actions, follow directly from the fact that the intention-in-action is formed and acted upon while an action tendency with control precedence is up and running. As a result, impulsive emotional actions manifest urgency, namely a preference for early versus late action, partial informational access, because the investment in information gathering and its quality are constrained by the pre-existence of a prioritized action tendency, and bodily underpinnings, because one of the elements of control precedence is bodily preparation.' (ibidem)

Furthermore, the flexibility element is also explained by how a prioritised action tendency only determines an abstractly-described relational goal, allowing compatibility and executive control to then fill in the details.

I will now briefly outline how MTE deals with the challenges of intentionality and differentiation. First, he also endorses Prinz's teleosemantic framework of intentionality, according to which emotions represent in virtue of their reliable causation. However, Scarantino introduces the function of achieving a certain relational goal, namely an informational-cum-motivational one (the function of both informing and motivating). So, MTE argues that 'what explains intentionality is that they are (in)action tendencies or action reflexes with the informational-cum-motivational function of achieving relational goals while correlating with core relational themes' (p.178).

When it comes to differentiating between different emotion types, MTE employs three different criteria. First, there are emotion-specific (in)action tendencies or action reflexes, such as attacking in the case of anger or disappearing in the case of shame. Secondly, there are emotion-specific core relational themes, such as offense for anger, or failure to live up to an ego ideal for shame. Thirdly, there are emotion-specific

relational goals, such as making up for a flawed behaviour for shame, or one's own safety for fear. These may, of course, sometimes overlap at the level of bodily changes, but there is a clear way for MTE to answer the challenge of differentiation.

6. Recalcitrance

In this section I will firstly define recalcitrance and then explore the different ways in which the accounts of the emotions discussed thus far deal with this challenge.

Firstly, D'Arms and Jacobson (2003) define recalcitrance as follows: 'We will say that an emotion is recalcitrant when it exists despite the agent's making a judgment that is in tension with it' (p.129). As it will have transpired from the discussion up to this point, the topic of recalcitrance gained relevance in virtue of being a significant issue for judgmentalist accounts. To reiterate, if an emotion essentially consists in a judgment such as 'The spider is dangerous' (fear), then it becomes problematic to explain the phenomenon of recalcitrance. As D'Arms and Jacobson point out, if you believe that something is not dangerous, you consequently (or so long as you entertain the judgment) cannot be afraid of that thing. So, recalcitrance would be impossible on this view. But, of course, this is obviously false. They go on to argue that judgmentalists should not infer a belief in something's dangerousness merely because one is afraid of it, for that would make their claim about the necessity of belief into a tautology, as it would always be true that there is a belief that p whenever there is an emotion. The more plausible way for judgmentalists to incorporate the possibility of recalcitrance would be to conclude that those experiencing recalcitrance are being irrational in virtue of holding opposing judgments at the same time (p.129).

This solution is less than ideal. For one, as D'Arms and Jacobson note, it seems very implausible that we would have such an irrational mental state, given that we are often acutely aware of our conflicting judgment (those afraid of flying know very well that it is safer than cars). So, could we consciously endorse both p and non-p at the same time? Furthermore, that would also make our behaviour incoherent, for, despite our fear of flying, we still purchase tickets, are not afraid for our friends when they fly, or do not buy insurance whenever we fly. Given that recalcitrance is then a

significant challenge for judgmentalism, some authors have tried to change the theory to account for recalcitrance. 'Quasi-judgmentalism' is what D'Arms and Jacobson call such views that attempt to lessen the cognitive requirements of judgmentalism, while maintaining the key tenet about its content (for example Solomon's view discussed in 3.1).

Specifically, such views loosen the requirement that one's (emotional) thoughts must be affirmed by the agent. All that is required is the presence of some type of evaluative thoughts that are sometimes 'held in mind by intentional states of comfort or discomfort'. The example they provide of this kind of view is Roberts' construal view, since it involves, as they put it 'propositional attitudes short of beliefs' (p.131). They then attempt to show that quasi-judgmentalist views still struggle to account for recalcitrance.

Their strongest argument against quasi-judgmentalism relies on the conception of 'cognitive sharpenings', which they take to be 'types [of emotion] constructed by specifying a subclass of instances of an emotion, or other affective state, in terms of some thought that they happen to share' (p.137). Examples of this would be homesickness, or even what they call 'tenure rage', an emotion one could stipulate, and which consists of all the instances of losing one's opportunity for a tenure position and feeling angry about it. They contrast these emotions with what they call 'natural emotional kinds', which are roughly the same as the group that I referred to as 'basic emotions' at the beginning of this chapter. They prefer 'natural kinds' over 'basic emotions' due to the disagreements about which emotions these are in particular, but broadly they are emotions such as anger, fear, disgust, envy, guilt, etc. (p.138). Further, 'natural kinds' are defined in virtue of their evolutionary functionality for them.

The first part of the argument is to show that there are differences in the role that thought plays in cognitive sharpenings as opposed to their role in natural emotion kinds. To that aim they look at some empirical claims about how emotions evolved for their adaptive value to help us deal with 'fundamental life tasks,' 'universal human

predicaments,' or 'recurrent adaptive situations'.¹⁴ Thus, they point out that one general view would be that natural emotion kinds are subserved by discrete, non-linguistic mechanisms, unlike cognitive sharpenings (p.139). Further, and essentially, it seems that recalcitrance is more common with natural emotion kinds than with sharpenings. Indeed, it seems absurd that someone would feel homesick despite being at home, or that they would experience tenure rage despite having obtained a tenure position. This is problematic for quasi-judgmentalism, since they would presumably argue that all emotions are like cognitive sharpenings, since they all rely on beliefs. They thus do not have the resources to explain this distinction between natural kinds and sharpenings, which seems quite plausible.

The second part of the argument relies on the idea of 'stable emotional recalcitrance', by which they mean 'a standing disposition to have recalcitrant bouts of a particular emotion' (p.140). That is, the phenomenon of having a recurrent fear of flying despite one's existing judgment about its safety, such that one would experience the same recalcitrance every time they fly. D'Arms and Jacobson note that quasi-judgmentalism may struggle to account for stable emotional recalcitrance: 'Why do the putatively constitutive thoughts of certain emotions continually reassert themselves despite their conflict with considered judgment?' (p.141). We (from a judgmentalist standpoint) may be prepared to accept occasional recalcitrance, yet stable recalcitrance does not make much sense on a judgmentalist picture. As they put it:

'Because the judgmentalist tradition is committed to defining the appraisals characteristic of emotional experience in terms of independently available concepts, it is forced to treat conflicts between an agent's emotions and her judgments as competing exercises of conceptual thought. (...) This makes emotional recalcitrance a strange sort of brute fact about certain concepts: that we tend to have recalcitrant, affect-laden thoughts involving them.' (p.142)

It thus seems that recalcitrance is a significant challenge for judgmentalism, such that being able to account for it would place competing views at an advantage. Feeling theories, as briefly hinted at, seem to have an easier time in answering this challenge, since they define emotions as essentially feelings, not judgments. We therefore avoid the unwanted charge of irrationality according to which we simultaneously hold two contradictory judgments. However, that is not to say that feeling theories do not have

¹⁴ They reference Ekman (1994), as well as Tooby and Cosmides (1990).

to provide an account of recalcitrance: after all, why do we feel like we are being in some sense irrational if emotions were just, say, a feeling of bodily modification? My body may react in different ways to different environmental factors, but that does not mean I am irrational for reacting inappropriately. So what explains our perception of ourselves as irrational? If I feel hungry after having just eaten (too much), I will just assume something is off with my body, or my self-awareness, but not that I am irrational. So, even if feeling theories seem to have an advantage, they still need to account for why we consider recalcitrance a mark of irrationality and strive to be rid of those emotions.

Of course, more sophisticated feeling views such as Prinz's, which account for the intentionality of emotions by means of reliable causation, may have an easier way out. They could say that the core relational theme exemplified by an emotion is what conflicts with our beliefs – however, more would need to be said to construct a compelling picture.

Perceptualist theories, on the other hand, may seem to be *prima facie* better geared to account for recalcitrance. As already stated in section 4, emotions can be seen as conflicting with judgment just as perceptual illusions are. That is, perceptual illusions are a form of conflict between the content of the two mental states, yet it is not a contradiction, and thus not straightforward irrationality. The issue with this solution, however, is that, like feeling views, it then become mysterious why we would regard recalcitrant emotions as irrational, but not recalcitrant perceptions. We have already seen in 4b) one possible explanation offered by Roberts. In Chapter 2 section 2.2 I look at another attempt at explaining this, advanced by Döring.

When it comes to motivational views, it seems that much less has been said regarding their way of accounting for recalcitrance (perhaps due to the fact that they are more recent). Deonna and Teroni argue, as highlighted at the end of section 5a), that it is not difficult to see how, on their view one would simply come to have a particular attitude about a state of facts despite knowing it to be otherwise. This discrepancy between attitudinal stance and judgment allows them to avoid the severe charge of irrationality faced by judgmentalism. However, since on their view the correctness conditions of

an emotion derive from whether the object exhibits the property we associate with that particular emotion, i.e. with that particular bodily stance, then there is also a clear way in which recalcitrance is irrational. This seems promising, however it may still struggle with the updated challenge of recalcitrance I propose in Chapter 2 (that between weak and strong recalcitrance).

When it comes to non-phenomenological motivational views such as Scarantino's, presumably he would rely on the informational-cum-motivational function of our emotional action tendencies according to his MTE. Like feeling views, he endorses a teleosemantic picture of intentionality. Thus, he may argue that recalcitrant episodes are irrational, since the informational-cum-motivational function misfires, yet not as irrational as a judgmentalist would have to say. So, like Deonna and Teroni's attitudinal view, he may have found the right balance to explain the irrationality charge of recalcitrance. However, it also remains to be seen whether it can deal with the challenge of strong and weak recalcitrance in the next chapter.

Thus, it seems that, while introduced as a significant challenge for judgmentalist views of the emotions, recalcitrance still remains a challenge for any view of the emotions. Regardless of the view espoused, one still needs to find a way to account for both i) when we are recalcitrant, we are not fully irrational in the sense of espousing contradictory beliefs and ii) we are still in some sense irrational, because we regard recalcitrance as a form of irrationality worth eliminating. As such, in the following chapter I will introduce a new way of understanding recalcitrance that will not only satisfy both i) and ii), but also provide a way to answer the challenge of recalcitrance regardless of the view of emotions espoused. If recalcitrance is merely an illusion, as I will go on to argue, then none of these views would have to offer an account of it. Particularly, my account will be of most use to judgmentalist accounts, which suffer most due to this challenge.

Chapter 2: Does (the conflict of) emotional recalcitrance exist? The challenge from strong and weak recalcitrance

1.1 Setting the stage

As we have seen in Chapter 1, the literature on emotions has a powerful test case to determine the strength of any one account: the issue of recalcitrance. To reiterate, emotional recalcitrance is the conflict between a judgment and an emotion. The first thing to note here is that, in order for any mental state to conflict with judgment, it seems to require some type of content that could be understood as conflicting with judgment. That does not imply the content needs to be propositional, but it needs to at least be evaluative in some sense (i.e. broadly pair the object with a value such as dangerous).

So, since recalcitrance seems to be an all-too-common experience, any view of emotions needs to account for it, and to do so within certain boundaries: if one regards emotions to be judgments, like judgmentalism, then the resulting state seems too irrational and exceptional or unusual (can we hold two contradictory beliefs at the same time?). However, if one regards emotions as something too far from judgment (e.g. perception), then the conflict would not seem rational enough (could a cough, a perception, a reflex, conflict with judgment?). Thinkers have been divided on this, some concluding that these emotions are irrational, while others have tried to challenge the nature of the conflict to preserve the intuition that these emotions may not be as irrational as one might think. I contend that the starting point is wrong: do we have good reason to think that recalcitrance, as defined above, really exists?¹⁵

Is there such a conflict in the first place, or is it merely apparent? Are there any reasons to doubt the existence of this conflict? I think we have several compelling reasons. First, the concept of recalcitrance rests on some tacit assumptions. Specifically, since the existence of recalcitrance is inferred from the (conscious) experience of conflict-like states, then integral to that experience is an assumption that the subject (at least

¹⁵ Of course, I do not argue that the experience of recalcitrance does not exist. Rather, I intend to investigate whether the conflict designated by this term really exists, or whether it is merely apparent.

roughly) is aware of the content of the conflicting mental states. After all, why would a subject keep telling themselves that the spider is harmless, if not because they assume they know what the content of the emotion is, i.e. that the spider is dangerous? So, the subject assumes they are (roughly) aware of the content of both states. It seems straightforward that in the case of beliefs we are likely aware of the nature of their content, i.e. I am usually not mistaken about my belief that p. However, awareness of the content of an emotion is far less straightforward. An emotion relies on both a formal and an intentional object (defined in 1.2), and the tacit assumption here is that we are aware of both: I supposedly know both *what* my emotion is about, and *what kind of evaluation* my emotion consists in.

One might here agree that we do not always know what our emotions are about, indeed certain emotions even require therapy to fully understand, but that is not the case for most of our regular emotions. So, since we are generally aware of the content of our emotions, why should we think that all recalcitrant emotions are emotions we do not understand? The reason is that recalcitrance involves persistence, *contra* the general understanding of it, and that persistence points to misunderstanding, as I shall argue. I will clarify this in what follows.

I contend that there is an ambiguity in how recalcitrance is defined, i.e. the state of conflict between one's better judgment and one's emotion. Particularly, there are conflicts between these states that are resolved immediately, with the emotions cooperatively responding to the beliefs: I am upset because I think that I got a low mark (say, 5 out of 10), but then I remember that the grading system for that particular task is different (5 is the maximum) and that I actually obtained the maximum mark. Upon analysing this conflict, my sadness disappears right away. It seems inappropriate to consider this emotion recalcitrant, even though I do have a belief state that conflicts with it, since it is *responsive* to it. To emphasise this point, compare this with beliefs: I acquire the belief that it is raining outside, but when I recall that actually there is a detailed hologram making it seem as if it is raining, my belief is revised. Both are instances of evidence-responsiveness (i.e. changing in order to reflect the evidence). If you think that the first case, that of the emotion being changed because

of the contradictory belief, is a case of recalcitrance, then would that mean that the belief in the second case is also recalcitrant? Surely not. Even though in the belief case there are two mental states that conflict with each other, they are not both conscious (or transparent to the subject) at the same time, so we cannot conclude that it is a case of recalcitrance. It follows that for recalcitrance, we need both mental states to be conscious or transparent.

However, how should this transparency be understood? On the one hand, we have seen that evidence responsiveness seems to be immediate: once both states become conscious, one of them shifts to cohere with the other one. In one case I discard a belief, and in the other I stop experiencing an emotion. There may, of course, be some delay, but only for the sake of disambiguating both states. I may recall the belief that there is a different grading system, but I need a few moments to infer what that means regarding my outcome. So, once we understand both mental states, we are ready to shift towards coherence. If, on the other hand, my emotion does not respond to the belief, it seems like it will persist for longer, as captured by the semantic origin of 'recalcitrance': it is strongly uncooperative and resists change. It will thus take more than becoming aware of (what we take to be) the conflicting belief in order to eliminate the recalcitrance. This is what I think the relevant understanding of recalcitrance is, which I dub strong (evidence-unresponsive) recalcitrance, as opposed to the type described above, which, as I will call it, is merely weak (evidence-responsive) recalcitrance. More technically:

Weak recalcitrance: C is a case of weak recalcitrance if and only if the subject of C has an emotion E and a belief B such that (i) E conflicts with B, and (ii) E and B do not persist after the subject becomes aware of the conflict.

Strong recalcitrance: C is a case of strong recalcitrance if and only if the subject of C has an emotion E and a belief B such that (i) E conflicts with B, and (ii) E and B persist despite awareness of the conflict between them two.

Note that I consider strong recalcitrance to be all there is to recalcitrance (i.e. weak recalcitrance, despite definitionally fitting under the concept of recalcitrance, is not actually what is at stake when discussing this phenomenon). I will therefore only refer

to strong recalcitrance when I use the term 'recalcitrance' henceforth. One might point out that I fail to take into account the voices in the literature that would (presumably) also not consider weak recalcitrance a type of recalcitrance. Indeed, an example would be Döring, who defines recalcitrance as an emotion that 'persists despite...' (2015, p.1; 2014, p.1), thereby excluding weak recalcitrance from the definition by highlighting the temporal/persistent aspect of recalcitrance. However, the literature does not seem to offer a definite solution to this worry, since some definitions do mention the persistence of the conflict, while others do not. At the very least this indicates an uncertainty in the literature about whether the temporal element is relevant to the recalcitrance, so I deem the strong-weak distinction a useful one given the rather vague approaches to this topic. Furthermore, it seems that the majority does define recalcitrance merely in terms of the conflict, such as: Majeed (2020), Grzankowski (2017, 2020), D'Arms and Jacobson (2003, p.129), Brady (2007, p.1).

Given this distinction, then the question 'does recalcitrance exist' becomes 'does strong recalcitrance (or evidence-unresponsive emotions) exist?'. How would one go about answering this question? It seems to me that we must establish why these emotions are unresponsive, as well as clarifying what it is for them to be responsive. In other words, we must elucidate why emotions sometimes seem to resist conflicting beliefs. Once we establish why emotions are resistant, we will likely have an answer to whether (strong) recalcitrance exists or not. Indeed, I will argue that it does not. To that aim, I will assess several views from the literature to establish both whether they are satisfying accounts of recalcitrance and whether they have the resources to explain its unresponsive nature. Generally, it seems that there are two broad ways of accounting for recalcitrance and its purported irrationality:

1. Concede that it is irrational, and explain how it is so (perhaps by showing that there is something about its biology that makes it such);
2. Deny that it is irrational by casting doubt on the nature of the conflict involved (for instance regarding emotions as not *asserting* that the world is a certain way, but rather embodying a different attitude which would thus be incommensurable with our beliefs)

My account would consist in a third option:

3. Deny that the emotion is irrational, but on the ground that the content of the emotion is misunderstood and so the conflict does not exist at all.

The conflict is actually between our belief and *our assumption about what the emotion means*: once we elucidate the content of the emotion, it is either consistent with our beliefs or otherwise evidence-responsive and thus changeable through other, more relevant beliefs. To anticipate, what we misunderstand is primarily the (intentional) object of the emotion. The structure of this chapter is the following: in 1.2 I further clarify certain key concepts, assumptions, and introduce relevant vocabulary; in section 2, I will focus on addressing and rejecting existing views of recalcitrance, specifically Majeed's in 2.1 and Döring's and Deonna and Teroni's in 2.2; then I will introduce my own account in section 3. If my response to the problem at hand succeeds, then any view of the emotions that can accommodate my account thereby does not need to explain recalcitrance, since it will possess the tools to explain why there is none.

1.2 Relevant terminology and key concepts

First I ought to clarify what I understand by 'changing one's emotional state', and by extension what it means for an emotion to resist beliefs. This is crucial, for there is a simple sense in which I can change my emotions: I can simply distract myself. Indeed, the empirical literature shows how shifting one's attention is a powerful tool in changing one's emotional state.¹⁶ I could remain sad about my low mark, or I could focus on the party I will attend tonight and thereby cheer myself up. This is not the sense in which I understand changing one's emotional state in connection to recalcitrance. What I understand instead is an attempt to change the emotion by specifically addressing (what we take to be) its content. If I am sad about a low mark, for example, I may tell myself that that one mark means nothing, it is just a number etc. That is, we address the content of the emotion with *semantically related beliefs*. If it

¹⁶ See Ochsner and Gross (2005).

works, then we have weak recalcitrance. If it does not, it seems that we have strong recalcitrance.

One might here think that the emotion could eventually vanish on its own, or that I may indirectly dissolve it by, say, attempting to change the state of facts, such as threatening the assessors in order to receive a higher mark (this would be loosely connected to its content). In the first case, the emotion is not 'resolved'; rather, our minds shifted attention to other things, our bodily states calmed down, but, importantly, if I later thought about that test again, then my emotion is likely to spring up again. As such, the actual content of the emotion has not been addressed, but its occurrence has been overwritten by other needs. I thus do not consider this a case of successfully changing one's emotion.

The second scenario is more complex. If my initial emotion is really connected with *having received* a low mark, then artificially changing the mark does not change the fact that I may deserve that low mark, even if it has been revised. As such, I may replace my sadness with the joy of success in eliminating the traces of that mark, but if I were to earnestly think about the initial object of the emotion, i.e. having my work being assessed as unsatisfactory, then my emotion will spring up again. So, this is also not what I understand by changing or addressing one's emotion. When inquiring how it is that we sometimes seem to have strong recalcitrance and fail to change the emotion, I mean how come it cannot be changed in the same way as the cases of weak recalcitrance, i.e. by reinforcing certain conflicting beliefs such that I stop experiencing the emotion.

Before moving on to the next sections and addressing views of recalcitrance, there are two concepts that are crucial for my discussion: emotions are generally taken to have both formal and intentional objects, as briefly discussed in Chapter 1. First, intentional objects are what the emotion is about. If I am angry at my partner, the intentional object of my anger is my partner. Second, formal objects are the properties emotions attribute to their intentional objects. For instance, if I am afraid of a dog, my fear conveys that the dog is dangerous, just like my anger at my partner conveys that they have wronged me in some sense.

Essentially, formal objects are what unifies different instances of the same emotion: my fear of the wolf, of failing an exam, of losing my keys are all a type of fear because they convey the same kind of property about their objects. Not only that, but formal objects also carry normative weight: if fear is taken to attribute a certain property to a certain object, then the fear is appropriate just in case the object does have that property (Deonna & Teroni, 2012, p.41).¹⁷ For instance, if I am at a funeral and I feel joy about it (note the intentional object must be the funeral itself, not, say, something I remember), then my joy is not appropriate because it mischaracterises its object (as something joyful).¹⁸

Before further discussing the question at hand, I need to make explicit a certain assumption necessary for my project. Specifically, I need to assume that emotions are evaluative in some sense. In Deonna & Teroni's words, 'according to the idea that emotions are types of evaluations, having an emotion amounts to apprehending the object of the emotion in evaluative terms' (2012, p.40). That is to say, whenever we experience an emotion such as anger, we apprehend the object of our emotion (say a person who slapped you without a reason) as being unjust. That does not necessarily mean we form a conscious belief attributing that property to the object (although that is a possible view under an evaluationist assumption), but broadly that we somehow pair the object with a value.¹⁹

In other words, I need to assume that emotions do evaluate the world in some sense (i.e. attribute formal objects to it), and I need to do that because otherwise it becomes mysterious in what sense emotions conflict with judgments. If emotions are not saying anything about their object, i.e. if they have no formal objects, then how could they be in conflict with something that does attribute formal properties to objects, i.e.

¹⁷ This is a rather simplistic picture, yet at this point I primarily aim to show what formal objects are generally taken to mean. I will address emotional justification in Chapter 3.

¹⁸ Emotions are generally considered to have these two types of objects, but not every theory will endorse this. Some would argue that what unifies emotion types is a certain phenomenology or body-state, such as the feeling tradition (Scarantino and de Sousa, 2018).

¹⁹ Thus, arguably all the views discussed in chapter one may fall under the scope of such an evaluationist assumption, with perhaps the exception of early feeling views such as the James-Lange one. Later views such as Prinz' integrate the idea of 'core-relational-themes', which are essentially types of formal objects, even though manifested differently.

judgments?²⁰ Two reasons to endorse this are, as stated above, that they individuate emotions and specify the correctness conditions and justification of an emotion. While there are reasons to doubt either of these claims, I will here assume them for the sake of the argument.

However, I think the more interesting reason to assume an evaluationist picture of the emotions is the mere existence of recalcitrant episodes. Why do we experience these conflicts if the emotions do not, in some sense, *say something* about their object? When our perceptions contradict our judgments, we do not feel as troubled by it, we do not feel like we are being irrational, and lastly, we are more willing to accept and not try to alter our perceptual state. For instance, I am more willing to accept that the stick is straight and I perceive it as bent (under water), than I am to accept that I know something is harmless yet fear it.²¹ Likewise, we are more prone to regard people experiencing recalcitrant emotions as being irrational than we are when people are experiencing recalcitrant perceptions ('Why are you so afraid, you know flying is safer than driving' versus 'The stick is actually straight, but I agree it's difficult, perhaps impossible to see it'). So, I think the fact that we do think there is some type of rational conflict here, one that we even try to resolve, is a strong enough reason to want to take a closer look at this particular conflict and to therefore assume there must be some evaluative level to the emotion for the conflict to make sense at all.

²⁰ We should bear in mind, however, that I may not even need to make this assumption. According to Grzankowski (2020), it might be better to turn the structure of the discussion on its head and give primacy to the normative conflict, which can exist without us committing to any particular logical conflict. Briefly, in having contrasting attitudes we are automatically violating a norm (whether the norm of truth for judgments or that of fittingness for emotion). As such, the conflict already occurs at the normative level, which gives us a bit more freedom in how we choose to interpret the logical conflict we normally focus on when discussing recalcitrance.

²¹ Indeed, one might here insist on the idea that we are bothered by recalcitrant perceptions to the same extent, since we try intensely to change our perceptions in cases of illusions such as the hollow mask illusion, the black&blue or white&gold dress and so on. However, the relevant difference here is that we do not feel like we are being irrational if we fail to alter our perception, whereas we do so with our emotions. What is bothering us is not *that* we are experiencing the illusion, but that we could experience it differently if we tried. This is not the case with emotions, where arguably what is bothering us is the existence of what we take to be an inappropriate or wrong emotion.

2. Two different solutions

In this section I will address two ways of explaining recalcitrance and assess whether they manage to address the question of strong versus weak recalcitrance. These strategies correspond to the two main ways of dealing with recalcitrance and its irrationality highlighted in section 1.1. I will reject both and then offer a new account which will better deal with this challenge in section 3.

2.1 The problem is the emotion

An intuitive approach regarding this challenge (weak versus strong recalcitrance) is simple: we cannot determine *why* the emotion does not change in certain cases because it is an empirical matter. The solution, then, may have something to do with the way the emotion was generated.²² Particularly, Majeed (2020) discusses some concepts that are relevant here. The first important distinction he makes is that of Greenspan between groundless emotions, so emotions that are not grounded in evaluative judgments of their objects, and grounded emotions. Majeed points out that this corresponds to a theory of pathways of emotional generation, which divides them as follows (p.5):

1. Thalamus-to-amygdala
2. Thalamus-to-cortex-to-amygdala

The first one is a quick process that does not require conscious experience of the stimuli, while the second one is much slower and involves conscious experience of the stimuli. Groundless emotions in Greenspan's sense can be accounted for thanks to the first type of pathway. However, the question arises whether that implies absolutely no perception of the stimuli. That is, one might argue that we still have unconscious perception of these stimuli, which would serve the role of cognitive bases (i.e. the information derived from other mental states that one needs for an emotion to be possible; e.g. one needs to think there is a bear in front of oneself in order to experience fear of the bear). The first type of pathway need not give rise to any recalcitrance in

²² This is suggested by Döring as well, as I will outline at the end of 2.2.

the way we normally understand it: it is just a quick process that does not reveal any evaluation. This case is familiar to us: imagine somebody being presented with some stimuli and their body reacting before they even get a chance to think about what they are perceiving. That agent would often deny that they are feeling the emotion, even though they do experience some of the associated symptoms. That could be the conflict we associate with recalcitrance and, crucially, we would not consider this as irrational as the other cases of (grounded) recalcitrance.

Majeed uses this distinction to point out that recalcitrance is used as a blanket term for all such instances, despite there being evidence of different degrees of irrationality in these cases. For instance, we would not regard the person experiencing a (recalcitrant) groundless emotion just as irrational as somebody in a grounded case. While I agree that there should be a more nuanced approach to recalcitrance, I am not sure this distinction can help answer the question I have posed.

First, how could this distinction answer the challenge of strong and weak recalcitrance? The straightforward answer would be to say that the reason why we might fail to dispel recalcitrance (so, cases of strong recalcitrance) is that they are groundless cases. However, why should we think that this is the case? Firstly, what is there to guarantee that the way in which an emotion was developed determines how it can be changed? But even if we bite the bullet and accept that groundless emotions are all cases of strong recalcitrance, we still have a problem: they are not the *only* cases of strong recalcitrance (as Majeed himself would no doubt agree, since he argues that there can be instances of recalcitrance in both groundless and grounded cases).

Importantly, there are recalcitrant emotions that are clearly based on judgments. Imagine being in a group of people where somebody makes a subtle remark about you. You sense there is some offense there, but you need a moment to interpret their words to establish whether they are offensive (and subsequently feel anger). In other words, your emotion needs to engage with certain judgments to confirm whether this is an offence. Now, one of those beliefs may be that the person must be joking, thus obtaining recalcitrant anger. However, the crucial aspect is that this instance of recalcitrance is not groundless, since it engages with judgments. That is, there can be

emotions that are not groundless, yet recalcitrant. So, if Majeed's distinction were to solve our issue, we would be able to say that we cannot solve strong recalcitrance because they are cases of groundless emotions, whereas cases of weak recalcitrance are not. However, the distinction between groundless and grounded does not neatly map onto strong and weak recalcitrance, so it seems that his account fails to explain the difference.²³²⁴

This type of strategy, namely of dividing emotions into two classes and arguing that recalcitrance solely occurs in one of them, is similar to D'Arms and Jacobson's (2003) strategy mentioned in chapter 1, section 6. To reiterate, they divide emotions in 'natural emotional kinds', which are a group of basic emotions defined primarily by fulfilling their evolutionary function (such as fear, anger, etc.), and 'cognitive sharpenings', which are 'types [of emotion] constructed by specifying a subclass of instances of an emotion, or other affective state, in terms of some thought that they happen to share' (such as homesickness and tenure-rage) (p.137). Their point was that recalcitrance only occurs in cases of natural emotional kinds, since these emotions are not as thought-dependent as cognitive sharpenings (it would be odd to experience homesickness when at home, for instance). Then, they would presumably argue that weak recalcitrance concerns cognitive sharpenings, while strong recalcitrance concerns natural emotional kinds, which would persist because these emotions are not as thought-dependent. There are a few issues with this account.

First, it is not clear that these two categories really do exhibit different levels of thought-dependence. For instance, it would be just as absurd to be afraid of a bear that is clearly not there as feeling homesick when at home; clearly, then, both categories do essentially depend on some beliefs. Secondly, it is not that straightforwardly the case

²³ Of course, this is not a rejection of all possible empirical explanations. Rather, this is a rejection of one of the better suited explanations for our issue since, as Döring suggests in section 2.2, it intuitively makes sense to inquire whether there is something particular about the development of the emotion.

²⁴ Wilson interestingly suggests adding a further distinction within the thalamus-to-cortex-to-amygdala pathway, namely that between emotions processed through the (adaptive) unconscious and those through the conscious mind (2002, p.128). There might be a correlation between emotions processed through the unconscious and strong recalcitrance. However, since he acknowledges that even such unconsciously processed emotions can be made conscious, they also have the potential to be changed. This is consistent with the account I will provide in section 3 (indeed, it even seems to support it).

that cognitive sharpenings cannot be recalcitrant. Take somebody who experiences tenure rage yet believes that them not obtaining that tenure-track position was fair: there would be a conflict between their emotion and their belief, just not the one about whether they did get that position or not. Thirdly, even if we accepted this division, it is not clear that its implications for the strong and weak recalcitrance challenge will work.

Saying that strong recalcitrance only occurs within natural emotional kinds does not itself explain the persistence; these emotions may be less thought-dependent than the other category, yet they are not therefore entirely thought-independent. So, why would they persist? Furthermore, we can think of a plurality of scenarios where natural emotions respond to evidence: learning that the realistic bear coming towards you is just a well-crafted costume, learning that your partner's forgetfulness about your birthday is an act in order to surprise you with a thoughtful present, are all thoughts that can lead to the cessation of a natural emotional kind (fear and anger respectively). Why would it work with only some of them? Would we therefore need to further divide this category into natural emotional kinds that respond to thoughts and those that do not? That would take us back to square one. This type of solution is thus not helpful towards solving our challenge.

2.2 The problem is the relation between the emotion and judgment

I will now discuss another way of explaining recalcitrance and establish i) whether it is a good account and ii) whether it helps explain the distinction between strong and weak recalcitrance. The account I will address has been introduced by Döring, a perceptualist whose aim in defining recalcitrance is to show that we are not being irrational for experiencing such states. She defends that conclusion by arguing that recalcitrance is a conflict, but not a contradiction between the two mental states – as is the analogous case of perceptual illusions, where arguably there is a conflict between what we *perceive* to be the case and what we *know* to be the case, yet we would not take this as evidence of self-contradiction. It is important to note that she is not trying to provide an account of strong versus weak recalcitrance, so whether her account

deals with this question will be based on the inferences I draw from her account. However, she does mention the possibility of changing our emotions to match our judgments (which she calls *recalibration*), so I think her account has the prima facie potential to answer our question.

What, then, is Döring's account of recalcitrance? Her account differentiates between the judgment and the emotion by arguing that the emotion is experienced in a different *mode of presentation*. Specifically, she notes how emotions are experienced in what she calls a 'default' mode. The default mode is the state in which we immediately perceive something as true in the moment – and we do that by focusing on those salient features that present the object as (e.g.) fearful. By salient features she means those features of the objects that together can be taken to represent the property we associate with the emotion we are experiencing. So, when it comes to a scary gorilla, it would be its aggressive eyes, piercing teeth etc. that we are perceiving (2014, p.10).

So, she argues that the emotion is perceiving the object in a certain way (as having a certain property), where perception is verisimilar (2009, p.245). That is to say, the emotion is merely presenting something as if it were true, whereas the judgment is directly asserting that something is true (i.e. that something indeed has a certain property). So, our emotional attribution of the property 'dangerous' to the gorilla does not *contradict* our judgment's assertion that the gorilla is not dangerous. However, they are still *conflicting*, since they are presenting reality in two different ways, yet the conflict is not a rational one, just like in the case of perception. She thus obtains her account of 'conflict without contradiction'.

However, there still is some tension between our better judgment and our emotion. At this stage she provides what could be her answer to the weak versus strong types of recalcitrance. She notes that in some instances we seem to be able to calibrate our emotions or perceptions. She uses the example of perceiving the train one is on as moving, when it is in fact still – one could recalibrate the perception in light of the knowledge that a nearby train is the one actually in motion, and thus come to see one's train as standing still (2009, p.246). She claims we are doing the same thing when we successfully bring an emotion to match our judgments. Here she ends with the

following observation: ‘if an emotion is revised in the light of better judgment and knowledge, this is not a matter of contradiction but of *calibration*’ (ibidem, my emphasis).

Thus, she concludes with a distinction between emotions that can be calibrated, and emotions that remain recalcitrant – supposedly weak and strong respectively. At this point she is aware that this merely introduces further questions, and it does not help explain the difference between weak and strong cases. However, she is happy to have merely provided a framework that could begin to answer these questions. Her final suggestion is that maybe what grounds the difference ‘is a matter solely of the malleability of the causal mechanisms underlying our emotions and perceptions’ (ibidem);²⁵ at the same time, she admits a more interesting answer would have to do with the emotions’ non-inferential content.

In other words, we *can* recalibrate our emotions, but we *do not have to*, since there is no contradiction. There may be a mismatch between the emotion and the judgment, just like there is one between the perception and judgment in cases of illusion, but that is not a problem. We can change this mismatch if we see fit (by calibrating). However, the cost of this conclusion may be too high: as indicated by ‘the Helm objection’, in eliminating the contradiction one also eliminates the conflict (Döring, 2009, p.245). The phenomenological claim we started from is that we experience a conflict that we perceive as a rational one; we think we are being irrational for being afraid of that spider or of flying, or for being angry at our partner when we know they did nothing wrong. Döring’s account, like perceptualism in general, isolates the emotions from the judgment to such a degree that it is hard to see how we could come to regard recalcitrance as an irrational conflict, while not doing the same for perceptual illusions.

Döring has a reply. According to her, this phenomenological claim is just a wrong intuition. She attempts to explain this intuition away by placing it elsewhere: it is not the emotion itself that is irrational, but rather the action it gives rise to. That action is

²⁵ I addressed this type of solution in 2.1, which I reject.

opposed to whatever action would be necessary for the reasoned pursuit of my goals (2014, p.5). However, I think this argument fails for the following reasons.

Firstly, I think a charitable interpretation of her claim is that the problem is the *impulse to action*, which can be present regardless of whether we act on it or not. That is to say, the moment I experience fear, I also feel the need to run. However, that need to run contradicts the fact that I know I am safe and should remain there. I thus think that the conflict Döring is here indicating is that between those two calls to action, one stemming from reason and the other from our emotions, fighting to take over. I may even be walking towards the bear now, but I still receive this impulse to run – and that is what is keeping me in a state of irrational conflict.

I cannot argue against the fact that this conflict does, indeed, arise. However, I disagree that that is the *central* conflict of recalcitrance. On the one hand, we can clearly distinguish between the two conflicts: one is a conflict regarding what I should *experience* (I should not be afraid), while the other is a conflict about what I should *do*. On the other hand, one could think of a plurality of cases where there is no clear call to action: either from the emotion, or reason, or both. Consider the case of ‘a sinking feeling’, or ‘feeling empty’; I could think I should be happy since nothing is wrong, yet still experience these forms of sadness. Neither the emotion, nor the judgment, would pull me towards any kind of action – the conflict is purely *about how I feel*. Consequently, I think her rejection of Helm’s intuition fails, since it cannot be based in our calls to action. As such, the main issue with her account still stands: it diminishes the rational quality of the conflict.

That issue aside, does Döring give us the right tools to differentiate between strong and weak recalcitrance? First, she makes an empirical suggestion to the effect that maybe some developmental mechanisms might be more malleable than others (I addressed this in 2.1). Alternatively, she hints that the solution might be more philosophically interesting. For the reason already mentioned (namely diminishing the rational quality of the conflict), I think her account may not be the most accurate explanation of recalcitrance. Not only that, but it does not help answer our question. Specifically, she argues that the cases in which we fail to recalibrate our emotions are

cases of recalcitrance. As such, she defines strong recalcitrance as failure to calibrate, and so weak recalcitrance as success of calibration. This, however, does not have any explanatory power on its own: why do we sometimes fail to calibrate? We have merely restated the problem with a different vocabulary.

Döring's is only one of the accounts that argue for conflict without contradiction. Other accounts, such as Helm's, argue along similar lines that emotions and judgments do different things (Döring, 2009, p.241). Helm's claim is more about the attitudes towards the property attributed to the object: while a judgment asserts that something is a certain way, an emotion merely assents to it being that way. In the end, however, I think all these accounts suffer from the same issue, which is that in their attempt to say that we are not irrational in cases of recalcitrance, they are not saying enough about how we manage to sometimes avoid recalcitrance (weak recalcitrance).

Before introducing my account, I will briefly return to a question I raised at the end of chapter 1, section 6. Specifically, it seemed that attitudinal and motivational views offer interesting accounts of recalcitrance, which fit under the overall strategy of this section. First, attitudinal views such as Deonna and Teroni's (2012) understand emotions to be attitudes (examples of other attitudes are beliefs, suppositions etc.), such that their formal objects are not a part of their content. Comparatively, the formal object of belief is truth, yet truth is not a part of any particular belief; likewise, the formal object of fear is dangerousness, but dangerousness is not part of the content of the fear. The fear is simply a particular feeling of action readiness. To restate briefly, the evaluative aspect of emotions is explained as follows: instances of action readiness are correct iff their object really does exhibit a certain property, e.g. the readiness to flee is only correct if its object really is dangerous. Thus, we can say that there is an intricate connection between action tendencies and the evaluative properties necessary for making those particular tendencies correct.

We can then understand recalcitrance as follows: our emotional attitude towards the object is wrong, since it does not satisfy the correctness conditions of that particular attitude (e.g. my tendency to run or kill the spider is not correct because the spider is not dangerous). The emotion is in that sense irrational, since it fails its normative

criterion. However, it is not as irrational as two conflicting judgments would be, simply because there cannot be a conflict of content: after all, the formal property is not a part of the content of the emotional attitude, it is merely its correctness condition. There is thus no contradiction of content. As already stated, this solution seems promising, however it remains to see how it deals with the challenge of strong and weak recalcitrance.

I think that their account cannot differentiate between strong and weak recalcitrance, at least not given its current parameters. Specifically, they do not have the resources to explain why some attitudes are highly responsive to evidence, while others are not. It would seem that all they can say is that some attitudes misfire, i.e. they do not satisfy their correctness conditions, while others do so – it is, after all, normal for any function of an organism to not work perfectly every single time, just like our beliefs or assumptions are sometimes wrong as well. While it is true that we are not perfect machines, we can often tell a story as to why our beliefs misfired: we misunderstood the evidence, we judged on insufficient evidence, etc. However, there seems to be no satisfactory story when it comes to emotional recalcitrance. We can imagine a scenario in which we initially get scared of a gorilla, but then see that it is inside a cage and our fear subsides thanks to the evidence. Why did the evidence work in this scenario, yet it fails in a plurality of similar ones?

What is more, recalcitrant episodes tend to repeat themselves, and our emotions consistently misfire in highly similar situations. I will, for instance, be scared of every single medium- or large-sized spider I see – how come my attitudes reliably misfires in the same type of situation? Say I live in a very dangerous area full of thieves and I naively form the belief that X is a nice person because they smiled at me. After I get fooled by people smiling at me a couple of times, I will simply stop forming that erroneous belief, I will regulate my attitude such that it stops misfiring. How come our emotional attitudes do not seem to be as (consistently) responsive?

So, it seems that Deonna and Teroni's account can at most say that strong recalcitrance simply happens because it happens. There is no way to differentiate between strong and weak instances. I will not go into detail regarding Scarantino's motivational view,

as it would seem to suffer from the same issue. Saying that the informational-cum-motivational function of the emotion misfires and thus fails its correctness condition, but is not thereby wholly irrational, is a promising response to the issue of recalcitrance. However, it would probably not have anything more to say regarding why it sometimes misfires and other times it does not, and how come it reliably misfires in similar circumstances. I thus conclude that these accounts also struggle to deal with the challenge of strong and weak recalcitrance, and I will proceed to outline my account.

3. My solution

My account of recalcitrance rests on a more elaborate understanding of the objects of our emotions. At the beginning of this chapter, I briefly outlined the concept of intentional object, i.e. whatever our emotion is about. However, these objects can be qualified in more detail. As Roberts points out, an emotion's intentional object (the gorilla I am afraid of, the friend I am upset with etc.) is bound with certain contextual factors. We are angry *at* somebody, *for* something, done *in a certain way*, and *in a certain setting* (2003, p.61). For instance, I am angry at my partner for an offense (making fun of me) in an intentional way and in front of my friends. So, when asked what we are angry about, Roberts points out we provide this whole story. He also notes that whichever the story might be, we sometimes focus on some elements more than on others (say sometimes I may focus on the offense, other times on the offender) (*ibidem*).

I would like to take Roberts' observations a step further and broaden the concept of intentional object, which I will do via two main claims. My first claim is that these types of contextual elements that we attach to the intentional object (whether we understand them as a part of the intentional object itself or something structurally related to it) need not all be actual: they could also be modal. So, for instance, the gorilla in front of me is caged. Not only that, let us say the cage is made of thick glass and it could not even reach its claws out towards me to attack or break it in any way (and I know this). However, I am still afraid of it. Since I think that the gorilla is not

going to harm me (it is not dangerous), my fear is recalcitrant. But am I therefore being irrational? That is yet to be shown.

According to the account I will now outline, what enables the recalcitrance is a misunderstanding of the intentional object. Specifically, the intentional object is more modal or counterfactual than we realise. First, the intentional object as normally understood is this gorilla in front of me. But which of its contextual factors am I focusing on? Currently, it is only standing there, doing nothing – so, what story is making me afraid? I argue that my fear is not about how the world is, but rather about *how it could be*. The world could have been such that somebody left the cage open, that the gorilla is stronger than we thought and breaks through the cage, that something hits the glass causing it to break and so on. These are all potential scenarios that could happen. So, I am not afraid of the gorilla because it is harming me, but rather because it *could* harm me.

Note that my claim is *about the modality of the intentional object*, not that of the formal object. That is, the emotion is predicating danger of the *possibility* of being attacked by the gorilla, so the formal object is the same as has been argued in the literature (dangerousness). What is different about my account, however, is that the intentional object encompasses a possible scenario, rather than merely the actual state of affairs. So, even if one may argue that dangerousness is already a modal property (i.e. it means that something *could* inflict harm), that does not affect my account, since I am arguing for the modality of the intentional object.

Consequently, we should rethink how we understand the content of the emotion that is purportedly in conflict with judgments. That is, the emotion is not simply saying that the gorilla is dangerous, but rather that *the possible scenario of the gorilla attacking me is dangerous* (for brevity, 'the gorilla could be dangerous', with the understanding that it is the intentional object that is modal, not the formal one). This also gives us an answer for why we stay recalcitrant: if our emotion attaches to a potential danger, telling ourselves that the object is not dangerous does not address the actual content of the emotion. It could still be the case that the gorilla is harmless in our context, but not if we also consider those counterfactual contexts. Likewise, saying that flying is

safer than driving does not contradict the content that a crash is possible (which would, of course, be dangerous).

As such, we can see why the belief would fail to change the emotion: in a way, the emotion is right, since it captures those possibilities that our reason seems to exclude. That may be for a good reason, namely that they are highly unlikely, yet that will not change how harmful those unlikely scenarios would be (they have, after all extremely high stakes – death). Interestingly, Nussbaum points out that the judgments we are so sure of in these cases are themselves merely estimates. Sure, we may think we believe that the gorilla is harmless, but more accurately we seldom have such certainties and may actually believe that the gorilla is *likely* harmless. This lends further credence to my point about the conflict of recalcitrance being merely apparent (2001, p.32).

There is, as mentioned at the beginning of this section, a second part to my account. Intentional objects are not limited by possible or modal scenarios – they may encompass even more considerations. Take the following scenario: your partner does not reply to your texts, and you get angry. You then tell yourself ‘Wait a minute, they’re at work. Of course they cannot reply’. However, you remain angry. In this kind of scenario, it seems that simply analysing the modal aspects of the intentional object (your partner could have not replied intentionally) does not suffice to justify the emotion. Instead, I will argue that the emotion is actually about an implication of these contextual factors and not the facts themselves. The facts being as they are, there would indeed be no reason to remain angry. However, as we have already seen, our emotions are not solely directed upon the immediately given facts.

Indeed, upon discovering that *p*, our minds begin considering reasons, explanations, implications, etc. Our minds thus do not focus solely on what I call ‘the core of the intentional object’ (here, the unresponsiveness). In this case, I think it is plausible that what the emotion is about is a potential implication such as: if my partner does not reply to my texts, it might be because they do not care about me enough. Once our anger is about them not caring about us, it is useless for us to reinforce the beliefs about texting, for, the fact that they cannot text you right now does not nullify the possibility that they might not care about you.

So, once the (fallacious) belief about them not texting back is added to one's cognitive base, additional considerations such as implications and justifications are co-activated, making it such that our emotions now hold onto more than just the initial state of facts (partner not replying): they also encompass what our minds contribute on their own (why, what else could they be doing, what does that say about me/us, etc.).²⁶ Note that both the initial fact (or the apparent intentional object) and the implication (or actual intentional object) share their formal property of being unjust (at least from the subject's perspective). This fact is crucial, for it allows us to mistake one for the other. If the (only) implication of them not texting was, say, that they must be buying you a nice present, that would not be unjust and hence not warrant anger. If it did not warrant anger, and assuming there is no other relevant implication, then the emotion would have to be about them not texting back. If that were the case, however, I think we would not have a case of recalcitrance. If the emotion was really about texting, and we told ourselves 'they cannot text back', then the emotion should cease accordingly.²⁷ In the cases in which it does not, I argue it is because our beliefs address the apparent intentional object, not the actual one.

One may here think that my account implies that identifying the 'correct' intentional object of the emotion results in cessation of the emotion. However, that is not necessarily the case. It may, of course, happen in some cases. Nevertheless, given that the emotion actually does fit its object (e.g. anger at the prospect of being cheated on), then it may behave like a non-recalcitrant emotion. Since it is not in conflict with a particular belief, it is under no pressure to change simply because it addresses a possibility rather than an actuality. For a parallel, when we experience emotions about

²⁶ One might here wonder whether there is a change of emotion or of intentional object, i.e. whether initially the emotion was about the absence of a text and then moved on to implications, or whether it was about the implications from the beginning. I think this will vary from case to case, sometimes the transition being instantaneous, and other times slower – however, this detail is not important for my claim, as all I need to establish is that by the time we experience the conflict of recalcitrance, the change will have already happened. That is, the emotion is already about the implications, regardless of how it transitioned to that point.

²⁷ One caveat here is that this may look differently depending on the circumstances. That is, changing one's emotion is a complex phenomenon that requires the cooperation of all the parts of the emotion, so delays may happen. For instance, one's feeling (e.g. the anger) might vanish, yet the bodily reactions (say, increased heartbeat) may persist. At any rate, whether this claim is accurate seems to be an empirical matter, one which I mention in the concluding section seems to be supported by empirical research, yet more research needs to be done.

fictional objects, being reminded that the characters are not real does not imply that our emotion must cease.

To summarise, my account postulates that, in experiencing recalcitrance, subjects are misunderstanding elements of their emotion because they are not always about actualities, but also about possibilities and implications. To put it visually, there is a web of interconnected thoughts, images, possibilities, considerations, implications, that gets activated together with the core of the intentional object – the emotion, then, may hold on to any of these elements, or perhaps even the set as a whole, such that we sometimes misinterpret which of these elements is central to the emotion (the core of the intentional object is not necessarily the centre of the emotion). In other words, one could argue that the content of an emotion is not p (where ' p ' stands for ' x is f '), but rather ' p^* ' (which stands for ' x^* is f ').

The conflict of recalcitrance is thus between our judgment ($\neg p$) and our interpretation of the emotion (p), and not the actual content of the emotion (p^*). We can thus conclude that we are not irrational in experiencing those emotions, since the emotion is not in contradiction with the judgment (p^* and $\neg p$); indeed, they are responding appropriately to their object (it will become clearer how in chapter 3). However, there still is a conflict because we rush and misinterpret that content, thereby failing to change the emotion (our intention prior to discovering the misunderstanding), which in turn makes us feel we are being irrational. This account also provides us with an explanation of why we sometimes are strongly recalcitrant: we simply fail to change an emotion because we misunderstand it. I will expand more on this account in Chapter 3.

4. Concluding thoughts

In this chapter I have offered a new account of recalcitrance, one in which the conflict lies not between the emotion and the judgment, but between our misleading interpretation of the emotion and the judgment. This view is, as I have argued, preferable to the alternatives explored, one reason being that it better explains why we have a gap between strong and weak recalcitrance. The other views explored

cannot convincingly explain why this is the case. However, I must here note that I have only engaged with two types of answers, while there are numerous alternatives that I do not have the space to address here. Nevertheless, most voices in the literature solely focus on explaining why the conflict does not amount to contradiction, and they thus do not offer any psychological explanation of why recalcitrance persists in some cases and not others. As such, the issue of strong and weak recalcitrance is an interesting question worth further expanding on.

Furthermore, my account seems to be supported by some existing psychological accounts, which aim to ground a relationship between (covert) thoughts and emotions. One such example (though somewhat outdated) is psychoanalysis, where curing certain symptoms means better understanding what our emotions are *actually* about, as psychoanalysis discusses our pattern of sublimating and repressing certain thoughts. For a more recent development, however, we could look at CBT therapy, which seems to be effective according to recent studies (Tolin, 2010). This type of therapy instructs us to identify the patterns of thought associated with our unwanted emotions (such as anxiety). So, looking into the beliefs associated with our emotions and how we may sometimes compulsively focus on some worst-case scenarios is one way to change our emotional state, including recalcitrant states.

Importantly, some strands of CBT focus on eliminating unwanted emotions without judging them. That is, a patient is instructed to become aware of the emotion for what it is, without judging that it is wrong or mistaken in any way. The anxiety of an overly-anxious person may be appropriate given the subject's experiences; however, it may also become problematic and get in the way of accomplishing one's goals. Thus, therapy's goal to help relieve these emotions does not mean that the emotions themselves are wrong or irrational, but at the very least that they are unwanted for diverse reasons. As such, my account seems to be coherent with the empirical results in this field. This is also consistent with the discussion of rationality I conduct in chapter 3.

Finally, one might wonder what consequences my view has for the normativity of emotions. That is to say, when is an emotion appropriate or justified, since sometimes

they are directed at modal properties, or at implications and interpretations? My view could face the potential issue that one starts constantly considering, say, terrifying counterfactual scenarios and thus keeps experiencing a perpetual state of fear due to what they are imagining. Surely, that is not justified. But how can we avoid this conclusion while also maintaining that there is something that the emotion is getting right, since fear at a terrifying possibility does seem like an appropriate response? I will address this question at length in the third chapter of my thesis (in section 3), by making reference to conceptions of appropriateness as fittingness, warrant, coherence, as well as strategic norms.

Chapter 3: The theory defended

We have seen in the previous chapter that the conflict between emotions and judgments considered to be the core of emotional recalcitrance is only apparent. That is, the conflict occurs between our judgment and our assumption or misguided interpretation of the content of the emotion. Thus, the emotion is not irrational in the sense of conflicting or contradicting our judgments. Indeed, given their intentional objects (the prospect of danger, or of not being loved), the reactions seem appropriate. However, we must carefully consider the limitations of that conclusion, to avoid licensing a host of intuitively irrational emotions. Therefore, in this chapter I will discuss possible issues with my account, which I will dispel by outlining a clear picture of emotional justification. That is, I will establish what it is that makes an emotion justified. First, I will clarify what the challenges are.

Before delving into the rationality issues, I will address a more immediate worry: that of introspection. Since my main claim is that recalcitrance is due to misunderstanding the content of our emotion, one might wonder whether that implies we are generally unable to understand our emotions. That would be a problematic implication, since it is questionable whether we really get things wrong so frequently (recalcitrance is quite frequent, after all). Not only that, but some recalcitrant episodes occur repeatedly: I am not only afraid of the harmless spider now, but virtually every single time I encounter a spider that is at least medium-sized. So, one would think that, given the repetitive aspects of some of these emotions, we would eventually come to understand their actual content. Yet, most instances of recalcitrance about the same type of object feel almost identical; of course, after years of being afraid of tiny spiders I may grow frustrated at my supposed irrationality, but I do not understand my state any better than the first time I was afraid of a spider. So, how is it that we are so utterly unable to identify the actual objects of our emotions? In other words, how does failure to introspect fit in with my account?

Another worry one might have concerns the extent to which emotions about possibilities, implications, etc. *can be* justified. Even if anger at the prospect of being

cheated on is intuitively acceptable since it pairs the evaluation of injustice with (the possibility of) an act that we would agree to be unjust, there is a sense in which one might want to argue it is unfair towards one's partner. The partner would understandably argue that this is an irrational emotion – since they did nothing wrong, why should they be the target of anger over a mere possibility? In other words, would my account imply that such 'eccentric' emotions are justifiable? Can we make sense of the idea that emotions about possibilities can be justified at all?

Finally, there is the worry of continuous emotions (mentioned at the end of chapter 2). Assume we have agreed that emotions about possibilities can be justified; do any limitations follow? How are we to fix the justifiability norms for such 'broad' emotions, i.e. emotions about possibilities, implications, etc.? I could for instance imagine an asteroid falling on me and dying, and subsequently start feeling scared as a result of my imagination. Firstly, there seems to be something right with this emotion, in that fear would be the appropriate reaction if this kind of scenario were real. Secondly, however, there seems to be a sense in which this must not be justified (we are here assuming I do not have any evidence it will happen), since that would allow me to constantly imagine terrifying scenarios and thus constantly be afraid by my own counterfactual scenarios. I could be in a perpetual state of fear, joy, anger, by constantly imagining scenarios that would give rise to such emotions – but surely that would be unjustified? I call this the 'worry of continuous emotions'. As such, I think my account needs to be supplemented by a clear set of justification criteria to avoid this unintuitive implication.

This chapter will therefore have the following structure. In section 1, I briefly go over the problem of introspection. Then, before embarking on an exploration of justification criteria for emotions, there will be an intermezzo in which I briefly defend the idea that emotions about possibilities can be justified by comparing it with ideas from the literature on fictional emotions. Once that is established, I discuss different criteria for rationality in section 3. Section 3.a) deals with cognitive rationality, while 3.b) deals with practical rationality. In outlining these criteria I also find and explore a solution to the worry of continuous emotions.

1. The worry of introspection

A worry one might have about my proposal is the following: does it not have false implications about our abilities to introspect? That is, how come we are fairly frequently unable to understand the actual content of our emotions, especially despite their tendency to occur repeatedly? Firstly, introspection is defined as ‘a means of learning about one’s own currently ongoing, or perhaps very recently past, mental states or processes’ (Schwitzgebel, 2019). So, I will first point out that most of the time we are not actually introspecting our emotions; we do not try to grasp more about our fear. Instead, we simply know we are afraid and proceed from that state of facts (i.e. we act on and think from the fact that we are afraid, without a need to question what the state is about; we run away, we try to calm ourselves down, etc.). As such, the fact that we do not always know the content of our emotions does not come as a surprise, since recognising *that* we are experiencing a certain emotion is generally the starting point of the subsequent acts and thoughts.

Additionally, even if one argues that we often do have the relevant knowledge about our emotions, that is consistent with my view. After all, we do manage to change our emotions in cases of weak recalcitrance, so their content is not entirely opaque to us. Of course, that is excluding cases where external factors influence whether or to what extent we manage (such as if something more emotionally powerful comes up, or when we are inebriated etc.). Furthermore, we are often unaware of a lot of the mental states underpinning our emotions, such as biases, trauma, or even other background beliefs we may hold. The fact that we may also be unaware of what exactly it is our emotion is doing is thus not an empirically problematic claim.

However, as with biases and trauma, we could also unveil some of these aspects of our mental states if we apply ourselves to it – indeed, psychological therapy often consists in finding out what our emotions are ‘really’ about (Wilson, 2002, pp.7-8, 121-122). Note that uncovering the real object of an emotion does not mean we have to change it – often the aim is to accept them and learn to work around them. Discovering one tends to be anxious about a situation S because of negative past experiences does

not mean that the emotion is wrong. Indeed, it is doing its best to avoid repetition of those negative past events. However, this often will get in the way of one's other goals, so therapeutic instruction tends to build on understanding what the emotion is doing to learn to expect it and work around it.

I think a parallel with self-deception and self-handicapping might provide an insight into why we are often unable to identify the actual object of the emotion (at least initially). That is, the actual content of our emotions may be in some sense harmful, and a self-deceiving mechanism could potentially protect us from that unnecessary harm. Going back to the texting example from chapter 2, I think it would be more harmful to entertain the thought that your partner does not care about you than to focus on them not replying to your texts. This is not eliminating the fact that there is a problem, i.e. there is something to be angry about, yet you focus on a smaller version of the problem in order to cope with it. Alternatively, empirical research argues that a lot of our mental content is nonconscious simply for reasons of efficiency (ibidem, pp.8-9), thus covering the cases where the actual emotional content cannot be obviously considered 'harmful'.

2. Intermezzo: Emotions about possibilities?

As mentioned at the beginning of this chapter, one might worry that emotions about possibilities are unjustifiable. If one's partner simply gets angry over the possibility of being cheated on when absolutely nothing happened, then something seems amiss, and not the kind of emotion we should consider to be potentially justified. Likewise, if I start imagining asteroids falling on us while in the middle of class, and start feeling afraid, there is a clear sense in which I am doing something wrong – in other words, there seems to be something 'eccentric' and unnecessary about these emotions. Are these types of emotions justifiable at all?

An interesting starting point would be to compare these scenarios with fictional emotions, i.e. emotions about fictional objects or states of affairs (e.g. pity for Anna Karenina, anger at Darth Vader), since there is a similar puzzle in the literature. The 'Paradox of emotional response to fictions' (Liao and Gendler, 2020) highlights how

one cannot have (genuine, ordinary) emotional responses to something one knows is not real, and yet we do have several emotional responses to fictional objects. That is, we are afraid at horror movies, we cry when watching *Titanic*, or are angry at a character in a novel – there is no question that we are emotionally involved in fiction. But are all these emotions irrational, simply because we know their objects are not real?

I do not wish to go into too much detail regarding this paradox, but I will highlight one answer that may be problematic for my account. Walton (1978) argues that fictional emotions are not full emotions. That is, he thinks they are what he calls ‘quasi-emotions’, since they do not behave like ‘full’ emotions. For instance, when we are watching a horror film, we do not go running out of the cinema the moment we see a monster, which is what we would do if we ever saw such creatures in real life. When the screen shows a murder being committed, we do not call the police, which we would do if we saw that same situation happen in real life. So, fictional emotions do not come with the same behaviour tendencies and reasons for actions as non-fictional emotions.

Likewise, if I am imagining my partner cheating on me, that should not give me a reason to shout at them, punish them, or question them on their behaviour and intentions. So, we notice a way in which both fictional emotions and emotions about possibilities deviate from ‘full emotions’. One may therefore infer that emotions about possibilities are also not full emotions, and that they therefore might not be liable to the same rationality criteria as regular emotions. One may even think that they are necessarily irrational, since they could be seen as implying that those possibilities are real, in the same sense in which a fictional emotion may imply that monsters in horror films are real.

Of course, there are also differences between fictional emotions and emotions about possibilities. For one, possibilities *could* come true (at least the kinds of possibilities we have been discussing so far). It may actually come to be that one’s partner cheats, or that asteroids fall on Earth – but *Anna Karenina* will never exist (perhaps fortunately

for her). So, it makes sense that we would take these possibilities a little more seriously when it comes to justification criteria.

For one, it seems absurd to say that someone going to war and considering the possibility of dying in battle is irrational for experiencing fear of that possibility. Indeed, the situation is not real, but a rational agent presumably should consider such possibilities before acting in order to make the best possible decisions. Interestingly, Williamson (2016) argues that being able to imagine a certain scenario or imagine oneself perform a certain action can be an important epistemic source ('means to knowledge', p.1). Whether I can imagine myself living in a house is relevant to whether I buy it or not. Likewise, when I imagine what it would be like to, say, go on a date with a person, feeling butterflies in my stomach is a good reason for asking them out. Thus, emotions about possibilities can be a reliable epistemic source, so we should not dismiss them as irrational simply in virtue of their objects not being actual.²⁸

Furthermore, even if one still believes that such emotions are not full emotions and therefore may have different normative criteria, that would not match our actual normative practices. Indeed, we seem to hold people accountable for the emotions they experience or do not experience as a response to hypothetical scenarios. For instance, imagine asking a bride-to-be to consider the possibility of her future spouse cheating on her, and she says she genuinely does not feel anything when she considers that possibility. 'Really?' one might inquire. 'You would not feel sad or angry?'. 'No,' she insists, 'even if I seriously imagine them kissing another person right now, I don't feel anything'. Would one not question her love for her spouse-to-be? Would one not think that she *should* feel anger or sadness? Alternatively, if you asked someone to imagine losing an important person in their life, and they experience joy at that prospect, would we not say their emotion is wrong or inappropriate? Thus, it seems

²⁸ This is consistent with a strand within the literature that argues that fictional emotions are not irrational because they play a crucial role in rational decision-making (Gendler and Kovakovich, 2005). Gendler and Kovakovich also discuss what they call 'simulated emotions', which are the emotions we undergo when we imaginatively engage with potential courses of action. That is, they are very similar with my emotions about possibilities, only they are actively used in rational decision making. Thus, there is a strong case for the justifiability of these emotions.

to me that, regardless of the type of emotion they are (whether full or quasi-emotions), emotions about possibilities are still subject to normative criteria such that they can be either justified or unjustified.

We thus have an answer to our issue: emotions about possibilities can be justified and are not irrational by default. To further clarify in what ways they can be either justified or unjustified, I will devote the next section to exploring normative criteria for the emotions and to addressing the worry of continuous emotions.

3. Emotional justification and the worry of continuous emotions

When it comes to emotional justification, a series of diverse criteria has been offered in the literature. Consequently, my aim in this section is to outline a broad picture of what it takes for an emotion to be justified, so that we can then apply it to the worries discussed in the previous section. My strategy for offering such a picture will be to first go over Scarantino and de Sousa's (2021) taxonomy of emotional rationality, which I will discuss in tandem with Majeed's (2020) norms for emotional justification. There is a broad overlap between the two taxonomies, yet there are also interesting differences.

According to Scarantino and de Sousa (2021), an initial distinction should be made between cognitive and practical rationality; cognitive rationality is concerned with the ways in which an emotion construes the world and its objects, whereas practical rationality looks at how emotions play a role in relation to our goals, regardless of how we define them. I will begin with cognitive rationality in section a), and then address practical rationality in section b).

a) Cognitive rationality

When it comes to cognitive rationality, Scarantino and de Sousa (2021) distinguish between three different types: rationality as fittingness, as warrant, and as coherence. Rationality as fittingness obtains just in case the object of our emotion truly has the property the emotion is attributing to it. For instance, if one is swimming with a shark,

then fear is a fitting reaction because sharks are dangerous. Majeed (2020) would presumably align this type of rationality with the principle of cognitive rationality, according to which an emotion is rational if it accurately represents the world it is trying to portray. When it comes to rationality as warrant, Scarantino and de Sousa (2021) emphasise that what matters is whether the object displays the relevant properties, not whether they are actually possessed. For instance, if one is swimming with a fake shark, then one would be warranted in experiencing fear, because the fake shark exhibits the property of dangerousness, even though it does not have it.

Lastly, Scarantino and de Sousa also discuss rationality as coherence under the category of cognitive rationality. Coherence demands that our emotion is coherent with our representations of what the world is like. Recalcitrance would be a typical example of an emotion that does not reach rationality as coherence – at least on the traditional understanding of recalcitrance. This type of rationality matches one of Majeed's (2020) principles for emotional justification, namely 'The consistency principle', which essentially states that we ought not to judge both p and $\neg p$.

Interestingly, Majeed has an additional principle that would roughly be aligned with the idea of rationality as coherence, namely 'Axiological rationality'. According to this principle, a subject is rational if the emotion fits some paradigm scenario, by which is meant 'the historic origins of an individual's experience of and capacity for the emotions involved'. Essentially, paradigm situations provide the objects of one's emotions and capacities for the emotions involved. That is, an emotion must also fit our personal histories in terms of development of our emotional dispositions. Say, if somebody is raised in a family of tamers, they would not develop the same fear reactions as most people in relation to tigers and lions, such that if at some point they do experience fear when facing a lion, that would be considered irrational (barring, of course, exceptional circumstances).

Having established some of the main types of cognitive rationality an emotion could be subjected to, I will now assess whether the emotions I discuss, namely the broad p^* emotions with complex intentional objects, satisfy these forms of rationality or principles.

First, there is rationality as fittingness. Herein lies an opportunity to obtain more clarity regarding my account: on the face of it, one might say that a recalcitrant emotion does not fit its object. Indeed, our belief seems to contribute to that conviction (flying is safe, you have no reason to be afraid). However, the intentional object is not merely the spider in front of us, or the person we are angry with: instead, it is about a web of mental states connected to that core. So, whether recalcitrant emotions are fitting depends on whether the complex intentional object (p^*) bears that property, not just the core. So, if we ask whether the possibility of being cheated on is unjust, or whether the possibility of being attacked by the currently engaged bear is dangerous, I would say that they are. I think that somebody who considers the possibility of losing their perfectly healthy parents is responding appropriately by feeling afraid or grieving. Yet one might disagree and push for a rigid understanding of fittingness, on which an emotion is only fitting if the object is actual or real.

However, fictional emotions, briefly discussed in section 2, are also arguably not fitting on such a rigid reading of fittingness, since they are, after all, not about real objects. But we do not consider fictional emotions irrational, specifically not unfitting.²⁹ If somebody is afraid of the monster in the horror movie, the fear fits its object. If one feels amused by a violent scene in which, say, a puppy gets hurt, we would consider that highly unfitting. I thus reject a rigid interpretation of fittingness conditions on which what matters is only the core of the intentional object, or ‘what is actually there’. What matters is what the emotion is about, regardless of whether it is real or present. Thus, I would argue that recalcitrant emotions fit their complex p^* objects (anger at the possibility of being cheated on, fear at the prospect of being robbed on the street). Assuming that fictional emotions are fitting, then so are the

²⁹ Some authors argue that they are always irrational because they imply believing in the existence of such fictional objects (Liao and Gendler, 2020). However, there has also been increased support for the idea that fictional emotions are not irrational, particularly thanks to their role in rational decision making (Gendler and Kovakovich, 2005). Due to space considerations, I will here assume this latter position. Nevertheless, I will also point out that it would be a dubious conclusion to think that fictional emotions are irrational, given how common they are in our day to day lives: we do not just tolerate them, we seek them out, which we arguably would not do if they were irrational. Furthermore, we do not consider each other irrational for experiencing such emotions; indeed, we even find it unusual for somebody to experience no emotion while watching particularly touching films.

emotions I have been discussing throughout my work. This, then, is the source of the intuition that the emotion does *some thing* right.

Then there is rationality as warrant, which is met if 'the object displays the relevant properties', as described above. Of course, one might here wonder what it is for something to 'display' properties such as dangerousness or injustice – surely one cannot simply look at a shark and spot dangerousness on its teeth. Thus, I argue that a more accurate way to understand rationality as warrant is the following: there is rationality as warrant if the subject perceives the object as bearing the relevant properties. That way, even if one has an illusion of a shark, one is still warranted in being afraid even if there is nothing that could possibly display the property of dangerousness.

When it comes to the broad emotions endorsed by my account, those alternative scenarios or implications included in p^* seem possible to the subject. One is afraid on a plane because a crash is not only theoretically possible, but it may be something they can imagine – its possibility is, in a sense, palpable. One is afraid of the possibility of not being loved by one's partner because that is a very possible and harrowing scenario. Consequently, these emotions are warranted for the agent; even if they are about alternative scenarios or implications, they are real possibilities for the subject, just like a fake shark looks like a real shark for an unknowing subject.

Lastly, rationality as coherence. As already established, recalcitrance does not fail to meet this standard of rationality, as the conflict of recalcitrance is actually between p^* and $\neg p$. Furthermore, as mentioned in Chapter 2 section 3, Nussbaum notes that the belief involved is also often not a $\neg p$, but rather something along the lines of 'unlikely that p ' (2001, p.32). We do not believe, for instance, that there is no chance that the plane would crash; it being safer than cars does not mean it is absolutely safe, but rather that the likelihood of it being unsafe is very small. Thus, not only is there no incoherence here according to my view, but if we accept Nussbaum's suggestion, the coherence becomes even more apparent: 'the plane is unlikely to crash' (previously $\neg p$) and 'there is a small chance that it will crash' (p^*), are very obviously consistent. I

will not address axiological rationality here, as it seems to be subject-dependent and not make a difference to our current discussion.

b) Practical rationality

According to Scarantino and de Sousa (2021), practical rationality (or strategic, as they call it) concerns the ability of an emotion 'to lead to actions that promote the agent's interests and properly relate to other processes that affect actions, notably decision-making'. Further, they distinguish two components of this type of rationality. On the one hand, an emotion is *instrumentally* rational insofar as it leads an agent to 'select means conducive to the agent's ends'.

On the other hand, it is *substantively* rational if it leads an agent 'to pursue ends that align with the agent's interests all things considered'. An example of instrumental irrationality would be panicking when diving because one's oxygen resources are running low. The panic makes the person forget that breathing heavily and swimming faster depletes the oxygen quicker, and they also forget the fastest route to reach the surface. They thus start to move frantically in all directions and completely deplete their oxygen. So, even though the emotion maintains the aim of safety, it chooses the wrong means towards it, as it exposes one to even more danger.

When it comes to substantive strategic irrationality, Scarantino and de Sousa note that it could be applied to both emotion types and tokens. For instance, grief could be argued to exhibit this type of irrationality, since it involves an unsatisfiable desire that the person were still alive, so it simply cannot lead to satisfying any of our aims. However, more commonly there are tokens of this type of irrationality, for instance getting angry during a job interview, which would negatively influence the outcome and thus stand against one's interest.

Majeed (2020) offers an alternative way to look at practical rationality, which he calls 'strategic rationality'. According to him, a subject exhibits strategic rationality if their emotion fulfils its function. Emotions are known to enhance our attention of certain features of our background; therefore, they fulfil their function if they make us focus

on the elements in our environment that are relevant to our goals (p.16). This highlights *how* our emotions can satisfy or fail to satisfy both instrumental and substantive strategic rationality – the diver is focusing on the lack of oxygen, instead of the most efficient way to reach safety, and the interviewee is focusing on the injustice rather than their goal. Thus we understand how our emotions may exhibit practical irrationality.

Lastly, Majeed also proposes the enkratic principle, which ‘requires of you (whatever person you are) that, if you believe at t that you yourself ought to ϕ , then you intend at t to ϕ ’ (p.13). Note that in order to meaningfully employ this principle in our discussion, we would need to endorse a theory of emotions on which they have motivational force, i.e. they incline us to act in certain ways. Here he provides Döring’s explanation that recalcitrance would fail to satisfy this criterion, since it would motivate us to act akratically. The example she provides is that, if we believe we ought to go hiking because it would be good for us, if we have a recalcitrant fear of heights, we would be inclined to act akratically due to that emotion (*ibidem*). In this work I have not taken a stance on emotional motivation, but if we do accept this principle for the emotions, it looks like recalcitrant emotions will be irrational from this standpoint, as will become clear in what follows.

So, how do the complex emotions I have been discussing fare from the point of view of practical rationality? Arguably, in the majority of cases they would not satisfy this criterion, especially in the continuous emotion case I introduced at the end of section 2. Even if fear may fit the prospect of an asteroid falling on oneself, and it may also match one’s aim at safety, it would fail to secure the right means towards it (not only is it difficult to defend oneself against such a situation, but it may arguably be impossible, depending on the size of the asteroid). Further, I would argue that such emotions also do get in the way of our interests all-things-considered, since they would have us focus on less likely situations. Sure, I may prepare myself by constantly keeping up to date with the asteroids passing by earth, but by looking at the sky and in the news, I fail to, say, keep up with local knife crime or thieves, or earthquakes and fires, which are more likely to affect me. Of course, that may also distract from more

immediate goals such as doing one's best at work, feeding oneself healthily etc. Thus, these emotions fail to secure both instrumental and substantive strategic rationality.

That is not to say, however, that they will go against practical irrationality in every single scenario. In the example with getting angry at one's partner despite learning that they did not actually cheat on the subject, the anger seems to convey or express how important or valuable the partner is to the subject. Indeed, by getting angry over the mere prospect of being cheated on can be seen as making the other person aware of the dangers of doing something of that sort. Of course, that may not be the best means towards that end, and it may backfire depending on the partner, but at the very least it does convey that information. Imagine a couple in an open relationship, and one partner casually tells the other that they were with another partner that day. If the person being told this gets angry, despite knowing they have no reason to given the terms of their relationship, that would inform both the subject and the partner that maybe an open relationship is not suitable for them. It may lead to a discussion of their current relationship status, and it could potentially lead to a better agreement.

Thus, we can say that in the majority of cases complex emotions (i.e. the emotions with broad intentional objects discussed in my account) fail to live up to practical rationality criteria, although there may be a couple of exceptions. When it comes to the dilemma we started from, namely that it does not seem that continuous emotions are fully irrational, yet there seems to be something irrational about them, we have found a solution: what these emotions are getting right is that they fit, cohere with, and are warranted by their object, yet they are not practically rational. Further, if one disagrees with my argument for the fittingness and warrant of such emotions by claiming that fittingness and warrant should only be about the core of an intentional object, rather than the whole package, my argument as a whole still stands because there still is rationality as coherence – so, the emotion is, indeed, getting some things right.

4. Conclusion

In this chapter, I have addressed potential worries one might have for my account. First, I showed that my account does not have implausible implications for introspection, as it does not imply we are generally unable to understand our emotions, but rather that we sometimes only assume their content. Second, I have defended the idea that emotions about possibilities, implications, etc. can be justified. Lastly, I have discussed the worry of constant emotions by establishing clear justification norms for emotions. The complex emotions discussed under my account do fit, cohere with, and are warranted by their objects, but they are not conducive to practical irrationality. Thus, we satisfy the intuition that continuous emotions are partly justified and partly unjustified.

Conclusion

This work aimed to challenge the idea that recalcitrant emotions are irrational by challenging the existence of the conflict supposed to be at its core. That is, I set out to explore whether the purported conflict between emotions and beliefs central to recalcitrance really exists – and I have shown that it does not. By focusing on the seeming evidence-unresponsiveness of these emotions, compared to the evidence-responsive nature of non-recalcitrant emotions, I developed an account that explains how recalcitrant emotions are not, after all, evidence-unresponsive. Instead, they cannot respond to the evidence provided simply because they are about a different object entirely. Thus, I have developed an account according to which the intentional object of our emotions, particularly recalcitrant ones, is broader than we might think: it includes a plurality of mental states related to the core of the intentional object, such as related possibilities, or implications etc.

By introducing the distinction between strong and weak recalcitrance, my discussion deviated from how recalcitrance has generally been approached in the literature. As we have seen in chapter 1, different views provide different answers to the challenge of recalcitrance. Instead of defining the emotions in a way that would satisfy the challenge of recalcitrance, this work looked at the necessary conditions for recalcitrance (evidence-unresponsiveness) in a way that tried to assume as little as possible about what the emotions are. Consequently, by deviating the discussion away from the literature and focusing on a new puzzle, i.e. the challenge from strong and weak recalcitrance, I have developed an account that can subsequently be integrated with most views of emotions.

In other words, my account is consistent with most of the existing views of emotions, and it is especially useful for judgmentalism, as it would help it avoid the unwanted conclusion that we are highly irrational whenever we experience recalcitrance. Not only that, but it nicely fits with existing views of emotional justification: emotions about possibilities can seem normatively ambiguous, as there are reasons why we might consider them both appropriate and inappropriate. However, I have shown

why that might be: these emotions fit, cohere with, and are warranted by their objects, but they may fail to satisfy criteria of practical rationality.

Bibliography

- Arkin, R. M., & Oleson, K. C. (1998). Self-handicapping. In J. M. Darley & J. Cooper (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 313–371). American Psychological Association. <https://doi.org/10.1037/10286-006>.
- Aristotle (1985). *Rhetoric*. Translated by: Rhys Roberts. In (eds.) Barnes: Complete Works of Aristotle. Volume 2. Princeton University Press.
- Barrett, L. F. (2006). Emotions as natural kinds? *Perspectives on psychological science*. 1. Pp.28-68.
- Brady, M. S. (2007). Recalcitrant Emotions and Visual Illusions. *American Philosophical Quarterly*. Vol. 44. No. 3. pp. 273-284.
- Cannon, W. B. (1927). The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *American Journal of Psychology* 39, 106-24.
- D'Arms, J. and Jacobson, D. (2000). The Moralistic Fallacy: On the 'Appropriateness' of Emotions. *Philosophy and Phenomenological Research*, 61(1): 65–90. doi:10.2307/2653403
- D'Arms, J. and Jacobson, D. (2003). The significance of recalcitrant emotion. *Royal Institute of Philosophy Supplements* 52. Pp.127-145.
- Deonna, J.A. & Teroni, F. (2012). *The Emotions. A philosophical introduction*. Routledge.
- Dow, J. (2009). Feeling Fantastic? - Emotions and Appearances in Aristotle. *Oxford Studies in Ancient Philosophy*. 37:143-175.
- Döring, S.A. (2009). The Logic of Emotional Experience: Noninferentiality and the Problem of Conflict Without Contradiction. *Emotion Review* Vol. 1, No. 3, pp.240–247.
- Döring, S.A. (2014). Why Recalcitrant Emotions Are Not Irrational. In: *Emotion and Value* eds. Roeser, S. and Todd, C.
- Downing, G. (2001). 'Emotion Theory Revisited'. In: *Heidegger, Coping, and Cognitive Science: a Festschrift for Hubert Dreyfus*. Vol. 2. M.I.T. Press. pp. 245-70.

- Ekman, P. (1994). 'All Emotions are Basic'. In Ekman, P. and Davidson, R. J. *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press.
- Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.
- Gallistel, C. R. 1980. *The Organization of Action: A New Synthesis*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Gendler, T. S. and Kovakovich, K. (2005). Genuine Rational Fictional Emotions. In: Kieran: 241–253.
- Grzankowski, A. (2020) Navigating recalcitrant emotions. *Journal of Philosophy*. ISSN 0022-362X. (In Press).
- Harris, R. N., & Snyder, C. R. (1986). The role of uncertain self-esteem in self-handicapping. *Journal of Personality and Social Psychology*, 51(2), 451–458.
- Liao, S. and Gendler, T. (2020). Imagination. In: *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2020/entries/imagination/>.
- Majeed, R. (2020). What Not to Make of Recalcitrant Emotions. *Erkennen*. <https://doi.org/10.1007/s10670-019-00216-0>.
- Nussbaum, M. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.
- Nussbaum, M. (2004). Emotions as judgments of value and importance. In Robert C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*. Oxford University Press.
- Ochsner, K. N. and Gross, J. J. (2005). The cognitive control of emotion. *TRENDS in Cognitive Sciences*. Vol.9. No.5.
- Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. New York: Harper and Row.
- Prinz, J. (2004a). Embodied Emotions. In Robert C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*. Oup Usa.
- Prinz, J. (2004b). *Gut Reactions: A Perceptual Theory of Emotion*. Oup Usa.

- Roberts, R. C. (1988). What an emotion is: A sketch. *Philosophical Review* 97 (April):183-209.
- Roberts, R.C. (2003). *Emotions. An Essay in Aid of Moral Psychology*. Cambridge University Press.
- Rossi, M. and Tappolet, C. (2019). What kind of evaluative states are emotions? The attitudinal theory vs. the perceptual theory of emotions. *Canadian Journal of Philosophy*. Vol. 49. No. 4. 544-563.
- Salmela, M. (2011). Can Emotion be Modelled on Perception? *Dialectica* 65 (1):1-29.
- Scarantino, A. (2014). The Motivational Theory of Emotions. In: D'Arms, J. and Jacobson, D. (eds). *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Oxford Scholarship Online.
- Scarantino, A. and de Sousa, R. (2021). "Emotion". In: *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/emotion/>.
- Schwitzgebel, E. (2019), "Introspection", *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/introspection/>.
- Solomon, R. C. (1977). The Rationality of Emotions. In: Solomon, R. C. (2003). *Not Passion's Slave: Emotions and Choice*. DOI: 10.1093/0195145496.001.0001.
- Solomon, R. C. (2003). I. Emotions, Thoughts and Feelings: What is a 'Cognitive Theory' of the Emotions and Does it Neglect Affectivity?. *Royal Institute of Philosophy*. Supplement 52:1-18.
- Tappolet, C. (2016). *Emotions, Value, and Agency*. Oxford University Press UK.
- Tolin, D.F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clin Psychol Rev*. 30(6):710-20. doi: 10.1016/j.cpr.2010.05.003.

- Tooby, J. and Cosmides, L. (1990). 'The Past Explains the Present: Emotional Adaptations and the Structure of Ancestral Environment'. *Ethology and Sociobiology* 11. pp. 375-424.
- Walton, K. L. (1978). "Fearing Fictions", *The Journal of Philosophy*, 75(1): 5-27.
doi:10.2307/2025831
- Wilson, T.D. (2004). *Strangers to Ourselves*. The Belknap Press of Harvard University Press.
- Williamson, T. (2016). *Knowing by Imagining*. In: Amy Kind (ed.), *Knowledge through Imagination*. OUP.