THE IMPORTANCE OF EMOTIONAL EPISODES IN THE ACQUISITION OF MORAL UNDERSTANDING

Lucía Ontiveros Llamas

PhD Thesis, University of Nottingham

Acknowledgements

I am very grateful for having had the opportunity to become a researcher in philosophy. I have learned so much during the five years that I have spent studying in England, and I will treasure them forever.

I would like to start by thanking my PhD supervisors Neil Sinclair and Chris Woodard. I will never forget their talent, mentorship, kindness, and support.

I would also like to thank the Philosophy PhD community at the University of Nottingham for their friendship and enthusiasm. Special thanks to Craig French for his generous advice.

Thanks to IPADE Business School, and to Carlos Mata for their sponsorship during my PhD studies. Most importantly, I would like to thank Dr. Alejandro Armenta, given that, without his guidance, my thesis project would have been incomplete.

I wanted to thank my mother and my father†, my sister Betty, my brother Fernando, and their families for their love and support. Thanks to Simon Thunder in particular for offering me his help throughout the most difficult times.

And I would like to dedicate this PhD thesis to my dad, who I know would be very happy that I finished writing it.

Contents

Introduction

In this thesis, I will defend the view that, although moral understanding can be acquired in multiple compatible ways, instances of moral understanding correctly acquired via emotional episodes seem to stand *higher* in the scale of moral understanding in contrast with instances of moral understanding correctly acquired non-emotionally.

In chapter I, I introduce the case of *Moral Mary* who has spent the first eighteen years of her life at a moral laboratory emotionally 'sedated'. Inside the laboratory, Mary was trained to acquire as many instances of moral understanding why-*p* (henceforth just 'moral understanding') of types of action labelled as 'right' and 'wrong' as possible. The aim of the *Moral Mary* experiment is to show that without the experience of emotional episodes, her acquisition of moral understanding inside the laboratory is *deficient*. This deficiency will be made clearer throughout the subsequent chapters of the thesis.

In the first section of chapter I, I develop Mary's story, endorse Simple-K reliabilism as an analysis of knowledge (Dretske 1989), and discuss the possibility of moral knowledge for realist and some anti-realist views (Audi 2019) (Jenkins 2015). In the second section, I endorse a reductionist view of moral understanding (Sliwa 2017). In other words, I am committed to the claim that that every instance of moral understanding is constituted by some degree of moral knowledge (Ibid: 530). In the same section, I also argue that Alison Hills' view of moral understanding as the capacity for *moral reasoning* fails to capture other ways in which moral understanding can be acquired (Hills 2009). Finally, in the third section I outline my account of moral understanding in light of the Moral Knowledge Account (Sliwa 2017), and elucidate the way in which it can be regarded as a kind of epistemic sentimentalism.

In chapter II, I argue that moral testimony—the most common way in which moral knowledge and moral understanding can be acquired— is problematic mainly for the reason that the understanding that it provides is *insufficient for acquiring the highest level of moral understanding*. In the first section, I introduce the most widely accepted pessimistic theory which finds moral testimony problematic (*moral pessimism*) (Nickel 2001, Hills 2009, McGrath 2011, 2019), but mainly focus on describing Guy Fletcher's view (Fletcher 2016). Fletcher has argued that moral testimony is problematic due to the fact that moral sentiments—which are intimately related to moral judgements—are at least difficult to form on the basis of pure, direct, testimony. In the same section I also describe Laura Callahan's Affect and Motivation view, which aims to provide novel reasons to hold moral pessimism (Callahan 2018). In the second section, I elucidate Daniel Wodak's optimistic argument for approving on the basis of moral and aesthetic testimony, in order to criticise Fletcher's pessimism, as well as moral pessimism (Wodak 2019). Thirdly, whilst I agree with Wodak's claim that deferring to the moral testimony of others can

provide us with reasons to morally approve or disapprove, I challenge his optimism by arguing that having reasons to approve or disapprove is not sufficient for possessing the highest degree of moral understanding.

In chapter III, my aim is to argue is that, unless they misfire, emotions can play a positive epistemic role. By 'epistemic role' I am referring to the production or modification of justified true belief and understanding. Concretely, I will defend the claim that emotions can be sources of salience (e.g., de Sousa 1987; Hookway 2008; Elgin 1996, 2008; Ben-Ze'ev 2010). In other words, emotions sometimes *direct our attention* to certain aspects of a given situation (e.g., when she encounters them, Mary's fear of spiders makes certain features of spiders salient, as well as the ways in which she can get rid of the spiders).

In the first section of this chapter, I elucidate the theory of emotions that I will be assuming throughout the thesis, and which falls under the category 'hybrid evaluative-feeling' (de Sousa and Scarantino 2021). In the second section, I describe in detail the ways in which emotions can play different putative and positive epistemic roles, and present some objections to these claims (Elgin 2008). In the third section, I discuss the Perceptual Theory of Emotions and the fact that when emotions act like perceptual experiences, the content that they provide is non-conceptual (Tappolet 2016). Finally, in the fourth section I discuss some ways in which emotions can be misleading (Goldie 2008) and irrational (e.g., Brady 2007, 2009).

In chapter IV, my aim is to argue that *attention* is a necessary condition to acquire understanding in general. I intend to defend this claim, to show that, since emotional episodes can in some cases focus our attention on the morally relevant features of certain actions, emotional episodes can lead to the acquisition of moral understanding.

In the first section, I will first provide the definition of attention that I will be assuming throughout the rest of the thesis, and discuss how attention relates to consciousness (i.e., self-awareness) (Brentano 1874). Secondly, I will describe four features of attention (selectivity, clarity, phenomenal character, controllability), in order to later focus on the importance that *selectivity* and *clarity* bear on the acquisition of understanding. Thirdly, I will outline how the involuntariness of attention (James 1890) will be relevant for showing that it is possible to acquire understanding without making an explicit effort to do so.

In the second section, I will first provide a brief historical account of some positive epistemic roles attributed to attention to support the claim that understanding requires attention. Secondly, I will argue that given the selectivity and clarity that it provides, attention is necessary for the acquisition of understanding. In the third section, I will first elucidate the way in which it has been argued that emotions can direct attention, since this mechanism will be crucial to describe the roles that some moral emotions can play in chapters V and VI. Secondly, I will also argue that although desires are mental

states that in principle can also direct our attention (Schroeder 2007) and increase our understanding, the fact that emotions *always involve a feeling component* makes them better candidates to explain the acquisition of at least two instances of moral understanding that I have in mind.

In chapter V, my aim is to argue that the highest level of moral understanding *necessarily* requires an *emotional acquaintance* with morally appraised actions. In the first section, I will begin by elucidating the view that moral understanding comes in degrees (Hills 2009: 103) (Sliwa 2017: 537, 548). I will also do this by referring back to the case of Moral Mary, and by describing what would constitute the different levels of the different putative instances of moral understanding. Secondly, I will point out that it is possible to acquire understanding of morally appraised actions in multiple compatible ways (e.g., via testimony, moral reasoning, emotional experiences, performing *prima facie* moral actions, via imaginings, epiphanies, contemplation of a work of art, etc.). Thirdly, I will emphasise that a thorough account of moral understanding should consider upstream (i.e., non-practical), downstream (i.e., practical) and combined (i.e., both practical and non-practical) instances of moral understanding.

In the third section, I will describe what virtuous emotional acquaintance consists in. I will understand virtue as the ability to '…recognize requirements which situations impose on one's behaviour' (McDowell 1979: 333), and rely on the Aristotelian model of the emotionally virtuous agent (*NE*, II, V: 1106b). Given that the emotionally virtuous agent experiences emotions '…at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way' virtue will be necessary for acquiring correct moral understanding of putative moral actions (Ibid: 1106b). Besides the valuable upstream and downstream moral understanding that emotions can allow us to gain due to their components, virtuous emotion is also epistemically privileged given the *perceptual awareness* that comes with it (Starkey 2008: 425). I will then describe the perceptual awareness of virtuous emotion, and argue that putatively moral actions are understood to their highest degree by *emotional acquaintance* (e.g., Mary fully understands the wrongness of lying once she experiences different episodes of *guilt* related to lying).

Finally, in chapter VI, I will argue that in some cases, certain emotional episodes can change—and thereby improve—our moral perspectives, and even our moral behaviour. By 'moral perspective', I am referring to *the epistemic standpoint whereby agents identify the morally salient features that ground the set of general moral beliefs that they implicitly endorse* (e.g., 'acts of charity are right', 'lying to your friends is wrong'). I argue that, since typically moral emotions (e.g., compassion and guilt) make certain features of moral actions striking or *salient*, their experiencing them can direct our attention to new or different morally relevant features, thereby producing a change in moral perspectives, which in turn can involve an improvement in our moral behaviour.

Firstly, I briefly introduce some views that define certain emotions as moral (Gibbard 1990) (Haidt 2003). I intend to argue that a certain experience of *typically moral emotions* can involve a change in

moral perspectives, given that these emotional episodes can constitute, arise from, or are associated to moral judgements (Prinz and Nichols 2010:112). Thirdly, through some examples, I argue that a shift in our second-order moral views —i.e., those views and/or attitudes about the appropriate grounds of moral judgements (Sinclair 2021:194)—brought about by an emotional episode can lead to a change in moral perspectives.

I will argue that even though inside the moral laboratory, Moral Mary possesses a high degree of moral understanding (e.g., it could even be argued that she possesses the six reasoning abilities involved in Hills' account of moral understanding to the greatest extent), she still lacks the information and evaluation provided by emotional episodes. I argue that she would also lack the distinctive motivation and feeling experience that they can cause, which in turn can all contribute to an increase in understanding of morally appraised types of action. I argue that these components, together with the attentional focus on the *prima facie* morally relevant aspects of an action or a situation that can be provided by emotions show the importance of emotional episodes in the acquisition of moral understanding. I also argued that virtuous emotional acquaintance with various morally appraised types of action consists in part in the *highest level of moral understanding*, and that some emotional episodes can lead to a *change in moral perspectives*. My Moral Epistemic Sentimentalism (MES) describes emotions in a positive epistemic light, but it is also cautious, as it acknowledges and discusses the ways in which emotions can misfire or hinder our acquisition of understanding in general.

Chapter I
What is Moral Knowledge? An Outline

In this first chapter, I intend to describe and defend the account of moral knowledge that I will be assuming throughout the rest of this thesis. In the first section, I will first introduce the case of *Moral Mary*, who for eighteen years has lived in a moral laboratory without experiencing any emotions. The aim of the experiment will be to compare Mary's acquisition of different kinds of moral knowledge and moral understanding inside and outside the laboratory, once she leaves and recovers her capacity to experience emotions. Ultimately, I will defend the claim that at least some emotional episodes will increase Mary's level of moral understanding despite her already having a great amount of moral knowledge, and despite her possession of moral reasoning abilities to a great extent. Second, I will explain my motivations for endorsing Simple-K reliabilism as an analysis of knowledge (e.g., Dretske 1989; Nozick 1981), and argue for the possibility of moral knowledge (e.g., Audi 2019; Shafer-Landau 2003). I briefly discuss the possibility of moral knowledge according to realism, and suggest that some anti-realist views can rely on an *explanationist* condition of knowledge to account for moral knowledge too (Jenkins 2015) (e.g., quasi-realists).

In the second section, I will first compare Alison Hills' nonreductionist account of moral understanding with Paulina Sliwa's reductionist account of moral understanding (Hills 2009, Sliwa 2017). Following Sliwa, I will argue that moral understanding is the ability to know right from wrong (Sliwa 2017: 523). In contrast, Hills argues that an agent has moral understanding-why if they possess six reasoning abilities that involve following and giving explanations (2009: 102-103). Second, I will discuss some criticisms to Hills' account of moral reasoning abilities which appear to be constitutive of moral understanding-why (Hills 2009: 102-103). I will then describe some results that show some ways in which Mary's moral understanding-why improved outside the laboratory, and some ways in which emotions negatively affected Mary's moral understanding, and partially analyse them in light of Hills' view. I will point out that, even if according to Hills' view, Mary inside the laboratory Mary had *full moral understanding*, it seems that by prioritising the capacity of moral reasoning, Hills is disregarding the significance that the capacity of experiencing emotions has in the acquisition of moral understanding-why, and moral understanding in general.

Lastly, in the third section I argue that moral knowledge is knowledge that employs moral concepts (e.g., knowing why an action is *prima facie* 'right' or 'wrong', knowing that an action is 'right' or 'wrong') (Sliwa 2017:546). I describe Sliwa's Moral Knowledge Account, according to which the capacity of moral understanding is the ability to acquire moral knowledge (Ibid:546). On this view, agents can exercise their capacity of moral understanding and thereby acquire moral knowledge via moral reasoning but also via other cognitive mechanisms such as perception, imagination, intuition, affective responses, etc. Agents can be said to possess moral understanding even if they are not able to

articulate it (Ibid: 547-548). Sliwa's Moral Knowledge Account will be more compatible with my view of acquisition of moral understanding rather than with Hills' account of reasoning abilities, given that it allows to decentralize moral reasoning as the main ability that leads to moral understanding. I briefly introduce my *Moral Epistemic Sentimentalism*, and suggest that it will be able to provide a more thorough account of the ways in which Mary's moral understanding outside the laboratory can improve, once she has recovered her capacity to experience emotions.

i. Moral Mary[1]

Moral Mary was born in a moral laboratory where she remained until she turned eighteen years old. Inside this laboratory, she was taught by her parents and some supposed moral experts to distinguish putative morally right actions from putative morally wrong actions through images of different social contexts (e.g., at school, at work, at parties, with family and friends etc.). They also used cards with the labels "right", "wrong", "permissible" and "impermissible", to teach her to identify these kinds of actions. Put simply, Mary learned that certain actions involved in daily social activities in the world outside the laboratory were usually categorised into two groups: *right* and *wrong*. She was taught that the actions that were categorised as 'right,' were those actions unanimously approved by everyone she worked with in the laboratory. In other words, 'right actions' were regarded as permissible, and the performance of this kind of action was encouraged (e.g., to share, to help). On the other hand, the actions that were categorised as 'wrong', were unanimously disapproved of. In other words, 'wrong actions' were regarded as impermissible, and the performance of these actions was discouraged (e.g., to physically harm others, to lie).

Whilst growing up, her parents and the supposed moral experts used to tell Mary stories such as fables (e.g., *Aesop's Fables*), and she would be asked to identify which character had acted 'wrongly' and which character had acted 'rightly'. Later on, Mary would be given more sophisticated stories about 'moral actions' to read and to discuss, such as Dostoyevsky's *Crime and Punishment*.

The most significant feature of Mary's experience in the moral laboratory, is that she remained *emotionally sedated*. In other words, inside the laboratory Mary remained unable to experience any kind of emotion, such as guilt or empathy. Mary did not experience any emotional episodes until she turned eighteen and left the laboratory.

[1] The case of 'Moral Mary' that I will develop in this thesis belongs to the category that Brian Scott Ballard has called 'Evaluative Mary Cases' (2020: 114). Other examples of Evaluative Mary Cases which are built upon the analogy with Jackson's Mary (1982), include 'Dud cases' (Ballard 2020), and Goldie's 'Irene' (2002).

Inside the moral laboratory, Mary studied the ethical theories formulated by the most prominent western authors, and she read about moral dilemmas and metaethical views. She was also shown some films about impermissible moral actions (e.g., about the Holocaust and other films about genocide). Since Mary was incapable of experiencing any emotions, she could not 'feel' that what she was watching on the screen was wrong.[2] She was merely told that the films portrayed 'morally wrong' actions, so she formulated negative moral judgements about some of the actions performed in these films. Her teachers also showed her sequences of images of both *prima facie* permissible and impermissible moral actions (e.g., someone finding a wallet on the pavement and taking it to the police station, someone stealing a car).

At the beginning of her training, Mary's teachers would write the words 'right' and 'wrong' over these images, but when Mary was in her early teenage years, they stopped writing these labels over new images, and they would ask her to label these images herself. This was a way in which they would test and correct her ability to recognise putatively morally right actions from putatively morally wrong actions (henceforth just 'right' and 'wrong' actions), and the same method was applied when showing her films. In her early years, Mary mainly deferred to the pure moral testimony of her teachers. Her teachers were vegetarian, so, for example, Mary believed that eating meat is wrong just because her teachers said so.

As a result of her training, it can be argued that Mary acquired knowledge-that. For example, she gained knowledge *that* it is right to take lost wallets to the police station, that it is wrong to steal someone else's car, etc. From this knowledge she could also inductively infer more general principles (e.g., that it is right to try to return lost possessions, that it is wrong to steal, etc.). Mary was also capable of saying why certain actions were either right or wrong according to the testimony from her parents (e.g., it is wrong to steal because you would not like it if someone else stole from you). Hence, it can be argued that Mary also acquired some knowledge-why.[3] She was also able to say why according to different normative theories certain actions were either right or wrong (e.g., it is wrong to steal because this state of affairs fails to maximise happiness, it is right to help others because it passes the test of the categorical imperative).

---

[2] For the sake of telling the story of an emotionless Moral Mary, I am assuming that watching films about genocides usually involves the experience of negative emotions that may signal or cause awareness that what we are observing seems normally *prima facie* morally wrong or impermissible. I am also presupposing a definition of emotion that describes emotions as necessarily having a feeling component (Goldie 2000; Döring 2007; Prinz 2004; Helm 2009; Ben-Ze'ev 2010; Tappolet 2016). I will discuss my account of emotion in more detail in chapter III of this thesis.

[3] I am aware that at least some reductionists about knowledge-*wh* (knowledge-when, who, why, where) argue that '*S* knows-wh' (e.g., "Mary knows why Bob is upset") is reducible to "there is a proposition *p* such that *S* knows that *p*, and *p* answers the indirect question of the wh-clause." So "Mary knows why Bob is upset" is equivalent to "there is a proposition *p* such that Mary knows that *p*, and *p* answers the question 'why is Bob upset?' (Brogaard 2009: 445). Similarly, in this case Mary's instances of knowledge *why* X is right or wrong (and other instances of knowledge-wh) could be reduced to some instances of her propositional knowledge (the same reduction could apply to instances of understanding-wh). However, my aim in this section is to elucidate the specific types of moral knowledge that Mary acquired inside the laboratory (i.e., knowledge-that, why, how), and to avoid confusion I am temporarily distinguishing all of these types of moral knowledge from moral understanding-why.

Another way in which Mary's ability to distinguish 'right from wrong' was tested, was by being asked to provide autonomous explanations of her moral verdicts. For example, after labelling the image of someone stealing a car as 'wrong', Mary was asked to provide the reasons *why* she believed that stealing a car is wrong. Mary would go on to say that she thought that stealing a car was *prima facie* wrong because it was unfair, and because it made more sense to her that everyone had a right to hold on to what they had worked hard to get, etc. This was at least one way in which Mary further developed the ability of moral reasoning (i.e., providing explanations). Mary's moral reasoning was evidence of her moral understanding-why.

Another way in which Mary developed her ability of moral reasoning was by drawing distinctions among certain putatively moral actions. For example, Mary was able to notice that it seems more important to be respectful to respectful people than it is to be respectful to rude people, even if it is generally important to be respectful, and that, in principle, it seems worse to lie to your mother than it is to lie to a stranger. She also became able to provide various examples of putatively moral actions of her own (similar but not identical to the ones that she memorised from the images and films); for example, that it is right to be supportive of others' accomplishments, or that it is wrong to reveal a secret that someone else has trusted you to keep. Additionally, besides asking Mary to discuss classic moral dilemmas in the academic literature (such as the Trolley Problem), the moral experts also encouraged her to discuss possible solutions to other kind of moral dilemmas; by asking her what she would do in certain cases, they thought Mary would acquire knowledge-how.

Hence, besides possessing a great amount of knowledge of putatively moral actions, Mary's reasoning abilities about these actions served as proof of her moral *understanding*. Succinctly, Mary's moral understanding inside the laboratory involved—at least in part—the following aspects and abilities. She was able to *reflect* about ethical theories, putatively moral actions, moral dilemmas, and metaethical views. She was able to *recognise* 'right' and 'wrong', 'permissible' and 'impermissible' actions. She was able *distinguish* 'right' actions from 'wrong' actions. She *developed* the *capacity of moral reasoning* by providing explanations, drawing distinctions, providing examples, and by imagining tentative answers to novel moral dilemmas.

The main aim of the hypothetical Moral Mary experiment is to help us consider the relationship between forms of moral understanding[4] and emotions.

[4] For simplicity's sake, for now I will treat knowledge and understanding as distinct mental states (see for e.g., Pritchard 2008; Hills 2009; Roush 2017). At present it is easier to describe the 'types of knowledge' -why, -that, -how, etc. that Mary has acquired and distinguishing them from the instances of understanding-why *p* that she has also acquired, without taking a stand on whether her moral understanding reduces to a sufficient amount of moral knowledge or not. Later on in this chapter, I will argue in detail for a reductionist account of moral understanding that defines moral understanding as knowing right from wrong (Sliwa 2017). Briefly put, I will argue that every possible instance of understanding -why, -that, -how, etc. is reducible to a sufficient amount of knowledge.

It is important to draw a distinction between the *capacity* for moral understanding and the outputs of this capacity. For example, Mary is capable of understanding certain distinctions between certain wrong actions; on the other hand, her understanding of why stealing a car is different from stealing a pencil would be a specific *instance* of moral understanding. Moral understanding *simpliciter* is the epistemic capacity to achieve instances of moral understanding, whereas instances of moral understanding are the mental states acquired by exercising the capacity. It is possible to have the epistemic faculty of moral understanding without achieving any instance of moral understanding (Sliwa 2017:524).

On her eighteenth birthday, Mary leaves the moral laboratory and starts to experience emotions. The moral experts confirmed that the first objective of the experiment was met, and it was proven that she acquired as much moral knowledge and moral understanding as possible. When Mary leaves the laboratory, she starts leading a conventional life and starts meeting people in different places. Presumably, Mary will continue to acquire moral knowledge and improving her moral understanding through her new life's experiences. She will also begin to perform some of the actions that she identified as right and wrong. However, the question underlying the objective of the experiment remains: in which way—if any— is Mary's moral understanding transformed by her new emotional experiences once she has left the moral laboratory? In chapter V, I will argue that Mary's new experience of emotional episodes can lead to an *increase* in her moral understanding of tokens of morally appraised types of actions, and in some cases to *the highest level* of moral understanding. In chapter VI, I will argue that after experiencing emotions outside the laboratory Mary can also undergo a change in moral perspectives, and that in some cases this change can lead to an improvement in moral understanding.

In the next subsection, I will briefly describe the conditions for knowledge that Mary's experiment presupposes. I will also discuss in a concise manner the kinds of metaethical theories that can accept the possibility of moral knowledge.

i.i What does Mary Know?

i.i.i General Account of Knowledge

My aim in this subsection is to lay the groundwork for an account of moral knowledge that a wide range of metaethical theories can accept. First, I will discuss Simple K-reliabilism, and my motivations for endorsing it. Second, I will provide a very basic description of moral realism and moral anti-realism, and argue that some anti-realist views can argue for the possibility of moral knowledge—for example—by adopting an expressivist strategy (Jenkins 2015). Third, I will argue that Simple-K reliabilism is an

analysis of knowledge that can be easily adopted by either realist or anti-realist views that accept the possibility of moral knowledge.

The analysis of knowledge that I assume in this thesis is Simple K-reliabilism (e.g., Dretske 1989; Armstrong 1973; Nozick 1981):

> *S* knows that *p* if, and only if, *S*'s belief that *p* (i) is true and (ii) was produced by a reliable cognitive process (in a way that degettierizes *S*'s belief).[5] (Steup 2001)

Nonetheless, any form of reliabilism would be compatible with the view of moral understanding that I will defend.[6] Analogous to the Justified True Belief Analysis of Knowledge, reliabilism about knowledge is a strategy to prevent lucky guesses from counting as knowledge (Ichikawa and Steup 2018: section 6). According to reliabilists, part of what is problematic about lucky guesses is that they are *formed* in such a way that it is unlikely that they should turn out to be true (Ibid: section 6). [7] In other words, reliabilism emphasises the importance that *cognitive processes* can have in the formation of true beliefs.

My main motivation for endorsing Simple K-reliabilism is that it can be easily accommodated in an Inclusive Project[8] in epistemology (Wild 2008: 125). Rather than being principally concerned with *warranting* true belief like the Traditional Project (i.e., views that typically hold the JTB account),[9] the Inclusive Project seeks to describe the various epistemic activities that can lead to the acquisition of knowledge and understanding (such as inquiry, belief formation or deliberation) (Ibid: 125). Moreover, the Inclusive Project incorporates not only beliefs but other mental states—such as emotions—which might contribute to the acquisition of knowledge and understanding (Ibid: 126). Describing the role that emotional episodes can play in the acquisition of moral understanding will be the main feature of my

[5] I am aware that there are complications in formulating cases that attempt to avoid the Gettier problem, but the clarification in parentheses is needed as a reminder that reliabilism faces the same problems as the JTB view. However, there are at least some cases in which our cognitive processes will prove to be reliable (e.g., sense experiences under normal conditions); thus, if our belief that *p* is true, and it is the case that *p* was produced by a reliable cognitive process, we will have knowledge that *p*.

[6] Another example would be Simple J-Reliabilism: Part A: *S* knows that *p* iff *S*'s belief that *p* is (i) true and (ii) justified. Part B: *S* is justified in believing that *p* iff *S*'s belief that *p* was produced by a reliable cognitive process (in a way that degettierizes *S*'s belief) (Swain 1981; Goldman: 1986).

[7] Although Causal Theories of Knowledge (Goldman 1967, 1976) do not concentrate on justification as a strategy to secure knowledge, I am only endorsing Simple K-reliabilism given that my aim in this thesis is to emphasise that the emotional direction of attention can be regarded as a reliable cognitive process involved in the acquisition of moral knowledge, rather than in the casual connection between our moral beliefs and the moral facts that they attempt to describe.

[8] See Elgin (1996), and Hookway (2003).

[9] See Stroud (2000), and Williams (2001).

view, so in that sense the claims that I will defend will be *more* compatible with the Inclusive Project than with the Traditional Project in epistemology.[10]

Hence, by endorsing Simple-K reliabilism my focus will not be in the justification of knowledge (either internal or external). Simply put, the theory of acquisition of moral understanding that I will develop will describe some emotional episodes as cognitive processes that can amount to knowledge and understanding in different ways. However, I will not be committed to the claim that emotional episodes can serve as justification for our evaluative beliefs.[11] Since the focus of this thesis is to highlight the importance of certain emotional episodes in the acquisition of moral knowledge and moral understanding, I will now proceed to discuss the possibility of moral knowledge.

i.i.ii The Possibility of Moral Knowledge

Accepting the possibility of moral knowledge is not an issue for moral realism. Roughly put, moral realists argue that moral facts and moral properties exist (e.g., the properties of rightness and wrongness). Moral realists argue that these moral facts and moral properties are mind-independent, and that the moral judgements that we make attempt to describe these facts (e.g., acts of kindness are *prima facie* right, acts of violence are *prima facie* wrong) (Shafer-Landau 2003:2). On the other hand, anti-realists hold that moral facts and moral properties do not exist, and that the judgements that we make about what is morally right and what is morally wrong are not descriptive; instead, moral judgements express a non-cognitive mental state of approval or disapproval (e.g., an attitude) (van Roojen 2004).

There are different types of moral realism (e.g., Intuitionism, Cornell realism), and different types of moral anti-realism (e.g., Emotivism, Quasi-realism). For instance, al classical *intuitionists* hold that basic moral propositions are self-evident, and that moral properties are non-natural properties (Stratton-Lake 2020: intro). According to Stratton-Lake, a *self-evident proposition* is 'one of which a clear intuition is sufficient justification for believing it, and for believing it on the basis of that intuition' (Ibid: 1.2, see Stratton-Lake 2016: 38). For Stratton-Lake, an *intuition* would be an *intellectual seeming* (as opposed to a belief, for example) (Ibid:1.2). So, Stratton-Lake says, what would justify our belief in

[10] Markus Wild argues that emotions can be integrated in the Traditional Project as 'ways of knowing' (e.g., in the general frame of virtue reliabilism), but as such they do not themselves contribute to the project of structuring knowledge and warranting belief (2008: 131). On the other hand, the supposed epistemic role of emotions is emphasised in the Inclusive Project, where emotions can play a double role: they can highlight the salient aspects of a situation and they can regulate epistemic activities (Elgin 1996, 2008; Hookway 2003, 2008). Although my view is similar to the Traditional Project in the sense that it will not hold that emotions can serve as justifications of our [moral] beliefs, it will become clear that the claims of my view resemble more the Inclusive Project where Elgin's and Hookway's view fits in. Wild has called Elgin's and Hookway's investigation 'Affective Epistemology' (2008: 127).

[11] See for example Johnston (2001), Döring (2007, 2014), Tappolet (2000, 2012, 2016). For a detailed defence of the view that emotional experiences involve perception of value see also Mitchell (2017). I will not be defending the claim that emotional experiences involve perceptions of value.

a self-evident proposition that *p* is that it seems true. For example, just like the answer to the question 'Why do you believe that the mug is red?' might be 'because it seems red,' the answer to the question 'Why do you believe that *p* is wrong'? might be 'because it seems wrong,' and what justifies both answers would be an intuition (i.e., an intellectual seeming that a certain belief is true). Furthermore, in brief, intuitionists in general argue that thin moral properties (e.g., goodness, rightness, badness, wrongness) are *non-natural* due to the fact that they cannot be defined wholly in terms of psychological, sociological, or biological properties (Ibid: 2.1).

In contrast, *Cornell Realism* is the view that moral facts can be investigated in a scientific way (Lutz and Lenman 2021: 3.2). Cornell Realism belongs to the category of *naturalism* among realist views. In other words, according to Cornell Realism, moral properties are like complex natural properties. For example, at least some Cornell Realists argue that the property of *goodness* is exactly like the property of *healthiness* in the sense that it is not directly observable, but nonetheless has a substantive causal profile (see Boyd 1988). Just like many things contribute to or detract from goodness (pleasure or pain, honesty or dishonesty), and many things can result from goodness in typical circumstances (human flourishing, political peace, etc.), there are many things that can cause or hinder health (nutrition or disease), there are many things that can result from health (energy, long life, etc.) (Ibid: 3.2). Hence, roughly put, according to at least a type of Cornell Realism moral properties are highly complex natural properties individuated by their causal profiles, and we know about them in the way that we gain scientific knowledge: via direct observation.

On the anti-realist side, for example, *Emotivism* holds that to make a moral assertion is to express an emotion, and emotions can move us to act in certain powerful ways, as well as to elicit similar emotions in others (Fisher 2014: 25; van Roojen 2018: 2.1). Assertions deploying general predicates of positive moral evaluation such as 'right', 'good', 'virtuous', etc., indicate a non-cognitive pro-attitude of approval, whereas assertions deploying general predicates of negative evaluation such as 'wrong', 'bad', 'vicious', etc., are indicative of negative non-cognitive attitudes (van Roojen 2018: 2.1). Hence, to call a person *vicious* is to express an attitude of disapproval, and the speech act of doing so is analogous to the speech act performed when we exclaim 'boo!' in reference to that person. As van Roojen explains, the same applies to predicative sentences deploying thick moral terms (e.g., cowardly, dishonest) which at the same time predicate natural properties (e.g., extreme fearsomeness in the case of being cowardly). In other words, thick moral terms can be understood as having both descriptive and emotive meaning (Ibid: 2.1).

Lastly, *Quasi-realists* are anti-realists who argue that we are entitled to act as if moral judgements are genuinely truth-apt, although strictly speaking they are neither true nor false in any robust sense (van Roojen 2018: 2.3). So, for example, to assert that 'bullfighting is wrong' can be a true statement in a deflationary sense. Simon Blackburn has argued that Quasi-realism is an attempt to explain (more an

explanatory programme than a position) our moral practice even though moral realism is false (Fisher 2014: 98). According to Quasi-realists such as Blackburn, although moral judgements express non-cognitive states (e.g., to say that bullfighting is wrong is to express an attitude of disapproval towards bullfighting), there is a need to make sense of some features of our moral practice, and so Quasi-realism is an ongoing project. For example, as Blackburn puts it: '… thoughts and concerns such as "I would like to know whether bullfighting is wrong," or "I believe that bullfighting is wrong, but I might be wrong about that," or "Bullfighting would be wrong whatever I or anyone else thought about it" – claims asserting our concern to get things right, our fallibility, and some independence of the ethical from what we actually feel.' (Blackburn 1993: 4)

I am aware that this is a very rough characterisation of realism and anti-realism, and for the sake of simplicity in the discussion I will not be addressing all of the possible distinctions among realist views, nor all of the possible distinctions among anti-realist views. I will not discuss the problems that all of these views might face. Since I will not argue for a specific metaethical theory, the theory of moral understanding that I will develop will aim to be compatible with either realist or anti-realist views which accept the possibility of moral knowledge.

An example of a realist account is Robert Audi's account of moral perception of rightness and wrongness. Audi argues that, for example, when observing that someone is slashing the tires of someone else's car, we are not only seeing the observable fact of slashing; our seeing also depends on what we already know, such as that the slasher does not own the car, as well as on our understanding of the normative significance of destroying someone else's property (Audi 2019:348).Additionally, Audi assumes that the wrongness perceived in the slasher's action is not a *brute* property of the slashing of tires, but rather that this action has the property of wrongness on the basis of or because of having other descriptive properties like, for example, being a transgression (Ibid: 349). Audi argues, 'Similarly, a person is not simply good but good *on the basis of* or *because of*, or *as* having, good governing motives together with beliefs appropriate to guide one toward constructive ends.' (Ibid: 349)

Hence, according to Audi's view, moral properties (e.g., rightness and wrongness) are not easily understood as being 'observable', however, they are apparently grounded on other properties[12] rather than on our beliefs and/or attitudes about them.

Note that Audi's account of moral perception would be compatible with Simple K-reliabilism if—for example— *S*'s belief that slashing someone else's car tires has the property of wrongness is true (*p*) and it was produced by a reliable cognitive process (in this case moral perception) (in a way that degettierizes *S*'s belief). Similarly, different realist accounts with different accounts of moral perception will render as reliable their version of the cognitive process at stake.

[12] Audi's theory of moral perception applies to both naturalistic and non-naturalistic moral realism (Ibid: 356).

Although all anti-realists deny the existence of moral facts and moral properties in a substantive way,[13] some of them can adopt expressivist views to account for moral knowledge (e.g., quasi-realists).[14] According to expressivism, our sincere moral judgements express moral attitudes (e.g., *prima facie* approval of acts of kindness) (Sinclair 2021:192). If the expressivist endorses the attitude of approval of acts of kindness (i.e., if they consider it appropriate), the difficulty would consist in explaining how can the expressivist *know* the proposition that acts of kindness are *prima facie* right. According to the expressivist view, our moral judgements do not express beliefs about a moral realm. Hence, for example, it is unclear in which way can they claim to know the proposition 'acts of kindness are right.' C.S.I Jenkins argues that expressivists can rely on an *explanationist* condition of knowledge (Jenkins 2015: 73):

> *S* knows that *p* iff *p* is a good explanation for an outsider[15] of *S*'s endorsement[16] of *p.*

In this case, *S* (the expressivist) knows that acts of kindness are *prima facie* right (*p*), if *S* can explain *why* they endorse *p* in terms of *p* itself. For example, *S* can explain that they *prima facie* approve of acts of kindness because these acts lead to peaceful interactions. A proposition such as *p* can feature in the explanation of *S*'s endorsement of *p* (i.e., *S*'s approval of *p*), as long as the expressivist denies that when *p* features in their explanation it is describing moral facts and moral properties. Rather, the function of *p* is to express an attitude that aims at coordinating our moral practice:

> Expressivists deny that moral judgments can be understood to express representational states that are answerable to the actual distribution of a distinct realm of moral properties (such as are believed to exist by the moral realist), but they insist that there is an alternative way of understanding moral predication and the standards governing it: the judgments are expressive of attitudes for the purposes of mutual co-ordination and the standards governing such expression are those appropriate to such a co-ordinating practice. (Sinclair 2012:155)

When the expressivist makes a moral claim such as "acts of kindness are right," they would be expressing their endorsement of an attitude of approval towards acts of kindness, and the grounds for their endorsement would be non-moral (in this case, *the fact that acts of kindness lead to peaceful*

[13] Moral error theories would be an example of anti-realist theories that could not account for moral knowledge (Mackie 1977) (Olson 2014). Standard versions of Error Theory argue that normative judgements (e.g., moral judgements) are *beliefs* that ascribe normative properties (e.g., moral properties), but that these properties do not exist and that all normative judgements are therefore false (Streumer and Wodak 2021: 254). For a characterisation of Error Theory that avoids formal objections (i.e., of the semantics and deontic logic kind) see Streumer and Wodak (2021).

[14] This epistemic strategy is not the only option that anti-realists can adopt to argue for an account of moral epistemology. (See Jenkins C.S.I., "What Quasi-realists Can Say About Knowledge", Oxford Scholarship Online, 2015: 72)

[15] This kind of outsider would be someone 'who knows nothing particular about *S* or *S*'s situation, but is rational, understands the content *p*, and knows commonplace things about people and their mental lives.' (Jenkins 2015: 72)

[16] The term 'endorsement' indicates that knowledge is also possible in situations where *S* has an attitude rather than a belief about *p*. (Ibid: 73)

interactions). Hence, as long as moral predicates such as 'right' or 'wrong' *refer to* non-moral grounds, and the expressivist's endorsement of certain attitudes of approval or disapproval *are caused by* non-moral grounds, then they can explain their use of moral predicates. *S* (the expressivist) is able to explain their endorsement of *p* (acts of kindness are right) to an outsider given that *p* (acts of kindness are right), since the reason why *S* approves of acts of kindness is because it leads to peaceful interactions. If it is true that acts of kindness lead to peaceful interactions, *S* can explain their endorsement of the claim that acts of kindness are right in terms of the rightness of acts of kindness.[17] Hence, the expressivist can meet Jenkins's explanationist condition of knowledge, and we can say that *S* knows that *p*.

Therefore, as long as anti-realists adopt a strategy such as the expressivist's, they can provide an account for their knowledge of propositions with moral predicates. Note as well that anti-realist accounts of moral knowledge which use the expressivist's strategy elucidated above are also compatible with Simple K-reliabilism, provided that *p* is true, and the explanation of the endorsement of *p* in terms of *p* specifies the exercise of a reliable cognitive process.

In this subsection I have shown that it is plausible to argue for the possibility of moral knowledge according to different kinds of realist accounts of moral knowledge, as well as anti-realist accounts of moral knowledge provided that these latter ones adopt certain epistemic strategies such as the expressivist. Also, both kinds of metaethical accounts also seem compatible with the analysis of knowledge held by Simple K-reliabilism, which fits better with an account of moral epistemology that focuses not on justification as a necessary condition for knowledge, but on other more relevant aspects such as our use of moral knowledge as well as on the cognitive processes of our acquisition of moral knowledge.

In the following section, I will compare two theories of moral understanding to further discuss what is involved in Mary's acquisition of moral knowledge and moral understanding.

ii. Moral Knowledge and Moral Understanding

In this section, I will compare two different theories of moral understanding in light of the case of Moral Mary. First, I will distinguish Hills' nonreductionist view of moral understanding from Sliwa's reductionism, in order to argue for a reductionist account of moral understanding (Hills 2009) (Sliwa 2017). Roughly, nonreductionists about understanding in general hold that it is possible to have understanding without knowledge (see Kvanvig 2003; Pritchard 2008; Strevens 2013; Hills 2016). On the other hand, reductionists claim that knowledge is necessary for achieving understanding (see Riaz

[17] According to expressivism, the explanation "S endorses the claim that acts of kindness are right given that acts of kindness are right" is equivalent to the explanation "S endorses the claim that acts of kindness are right given that acts of kindness lead to peaceful interactions." (See Sinclair and Chamberlain 2022: 59)

2015; Sliwa 2015). Second, I will argue that although inside the laboratory Mary possessed moral reasoning abilities to a great extent, she still did not possess the *highest degree* of moral understanding, due to her lack of emotional experiences.

ii.i Nonreductionism and Reductionism

Put simply, nonreductionists hold that moral knowledge and moral understanding are distinct. For example, according to nonreductionists, understanding-why lying is wrong is not the same as knowing-why lying is wrong (Sliwa 2017: 522). On the other hand, reductionists argue that when an agent understands why an action is right or wrong it is as a result of having knowledge about why it is right or wrong (Ibid: 525-526). I will briefly discuss nonreductionism first, in order to distinguish it from the reductionism that I will endorse.

Nonreductionists[18] specify some of their motivations for their approach. They argue, for instance, that it is easier to know something than it is to understand it, that to know something does not entail understanding it, and that knowledge can be easily acquired through testimony, unlike understanding (Ibid: 522). I will discuss, however, the case of moral testimony in detail in the next chapter.[19]

Nonreductionists such as Hills and Pritchard also argue that it is possible to have instances of understanding-why without having instances of knowledge-why as shown at least by some cases of epistemic luck (Hills 2009:104; Pritchard 2008:37).[20]

Hills gives the example of learning by chance that Stalin was an evil person:

> Suppose that your school has been sent a set of extremely inaccurate textbooks, which have been handed out to your class. But you are very lucky because there is only one that is accurate, and by chance you have it. You read in your book that Stalin was responsible for the deaths of millions of people. You draw the obvious conclusion that he was an evil person. It is plausible that you do not know that Stalin killed millions of people, since you could so easily have got a different textbook which was wholly unreliable. (Hills 2009:104)

In this case, Hills argues, you do not know why Stalin was evil because he killed millions of people, since you do not know that he killed millions of people (given that you could so easily have gotten a

[18] See for example, Kvanvig (2003); Strevens (2013); Zagzebski (2001).
[19] In chapter II I will discuss in particular Hills' pessimism about moral testimony. According to Hills, we can gain moral knowledge-that *p* and moral knowledge-why *p* by deferring to pure and impure moral testimony, but not moral understanding-why *p*.
[20] Some nonreductionists accept that instances of moral understanding require knowledge, but hold that knowing is not all there is to understanding (Sliwa 2017: 525).

different textbook which was wholly unreliable). However, Hills says that you can understand why he was evil, since your belief that he was evil because he killed millions of people is correct; also, we are assuming that you have both the ability to draw the conclusion that he was evil from the reasons why he was evil, and the ability to do the same in similar cases. Hence, this case of epistemic luck shows that it is possible to have moral understanding why *p*, without having knowledge why *p* (Ibid: 104).

Although it might be granted that in a case like this you do not know why Stalin is evil, it is still unclear why we should accept that, nonetheless, you understand why he is evil (Sliwa 2017:526). Sliwa argues that we can reject the premise that if you form a true belief why *p* by exercising the capacity of moral understanding, you thereby understand why *p*. If this were true, it would generalise to other epistemic faculties, and this does not seem to be the case. For example, forming a true belief that *p* based on exercising one's visual perception does not guarantee that one sees that *p*; you can see that there is a red apple in front of you based on your visual perception, but the object in front of you is an apple-shaped box rather than a real apple. However, there happens to be a red apple inside the box, so your belief that there is a red apple in front of you is true. You arrived at it by exercising your perceptual capacity, but it does not follow that you thereby see a red apple or that you see that there is a red apple in front of you (Ibid: 527).

In the same way, in Hills' Stalin example, you exercised your capacity of moral understanding and acquired a true belief, but this true belief is not an instance of moral understanding, and so you do not understand why Stalin is evil. Hence, as argued by Sliwa, Hills' argument from epistemic luck does not seem to show—at least in a case like this— that you can achieve an instance of understanding in the absence of knowledge, and so it does not provide us with a case to dismiss reductionism. According to reductionism, knowledge is all there is to understanding, and so every instance of understanding is constituted by an instance of knowledge (Ibid: 526). However, nonreductionists often characterise reductionism as:

> An agent understands why *p* if and only if she knows why *p*. (Ibid: 530)

However, reductionists are only committed to the claim that every instance of understanding is constituted by knowledge. These instances of understanding why *p* can be achieved depending both on whether agents know why *p* and on *how much* they know they know about why *p*. Consider Sliwa's reductionism:

> An agent understands why *p* if and only if she has a sufficient amount of knowledge why *p*. (Ibid: 530)

My main motivation for endorsing this kind of reductionism, is that it allows for many ways in which an agent can achieve instances of moral understanding why *p*, since there are many ways in which an agent can achieve moral knowledge why *p*. For example, via perception, intuition, moral testimony,

moral reasoning, emotional experiences, performing certain actions, via imaginings, epiphanies,[21] contemplation of a work of art, or through the creation of artistic works, etc. However, if moral understanding is defined as a concrete set of skills, there is a worry that multiple ways of acquiring moral understanding will not be identified, thereby limiting our inquiry of moral understanding acquisition.

In the next subsection, I will introduce Hills' nonreductionist account of moral understanding-why in order to continue to distinguish it from mine.

ii.ii Moral Understanding as the Capacity for Moral Reasoning

According to Hills, understanding why *p* involves a set of abilities (where *q* is why *p*) (2009: 102-103) to:

(i) follow an explanation of why *p* given by someone else;
(ii) explain why *p* in your own words;
(iii) draw the conclusion that *p* (or that probably *p*) from the information that *q*;
(iv) draw the conclusion that *p'* (or that probably *p'*) from the information that *q* (where *p'* and *q'* are similar to but not identical to *p* and *q*);
(v) given the information that *p*, give the right explanation, *q*;
(vi) given the information that *p*', give the right explanation, *q*'.

Hills argues that these abilities are necessary, and perhaps jointly sufficient for moral understanding why *p*, provided that it is true that *p*, and that *q* is why *p* (Ibid:103). On Hills' view, understanding why an action is morally right or wrong requires the ability to engage in moral reasoning about why it is right or wrong. Put simply, as shown by abilities i-vi above, moral reasoning encompasses the ability to give and follow explanations. Call this view of moral reasoning 'the moral reasoning claim' (Sliwa 2017: 523).

So, for example Mary's understanding of the wrongness of lying would require her being capable of learning that mutual trust is morally important relatively compared to other moral and nonmoral considerations (e.g., arrogance), as well as her being able to provide her own moral reasons not to lie. She should also be able to make appropriate judgements in similar cases, for example, that breaking a

[21] See Chappell, S.G., *Epiphanies: An Ethics of Experience*, Oxford University Press, (2022). According to Chappell, an epiphany is "an overwhelming existentially significant manifestation of value in experience, often sudden and surprising, which feeds the psyche, which feels like it 'comes from outside'—it is something given, relative to which I am a passive perceiver—which teaches us something new, which 'takes us out of ourselves', and to which there is a natural and correct response." (Chappell 2022: Intro i.i)

promise can be explained to be morally wrong for the same reason not to lie, but stealing is not (Ibid: 101-102).

Hills argues that if you are able to fulfil abilities i-vi, then you possess 'full understanding.' She says:

> You have minimal moral understanding if you correctly believe that q is why p and you can follow an explanation of why p. You have greater understanding the more you fulfil i–vi, and you have full understanding if you have i–vi to the greatest extent. (Ibid: 103)

Although Hills remains open to the possibility of there being other abilities required in addition to giving and following explanations, it nonetheless follows from her view that if Mary possessed abilities i-vi then her lack of emotional experiences would not represent a problem for her acquiring moral understanding-why (henceforth just 'moral understanding'). Hills says:

> 'Understanding is often associated with certain sorts of feeling: a flash of enlightenment; a light dawning. But these are not necessary: you need have no particular feelings at all when you finally come to understand why X is morally wrong. Nor are they sufficient. However much you feel the light dawning as you confidently turn aside from the needy, you do not understand—you cannot understand—that it is right to do so.' (Ibid: 103 footnote 19)

Hence, according to Hills, Mary would not need any emotion[22] to acquire full moral understanding. According to Hills' view, Mary's new emotional episodes outside the laboratory will not lead her to acquire moral understanding. In contrast, my aim in this thesis is to argue that emotional episodes can be distinctively useful in an epistemic way and lead us to acquire moral understanding. Once Mary has left the laboratory and she felt, for example, compassion or something like 'the light dawning' when *helping* the needy, then it is possible that she *would* at least in part understand that it is right to do so. However, if she felt compassion or the light dawning when turning aside from the needy thereby believing that it is right to do so, then such an emotional episode would misfire and lead her to acquire an *incorrect* instance of moral understanding.

On my view, Mary's moral understanding in this case would be incorrect since her belief 'turning aside from the needy is right' would be false, and—in accordance with Simple K-reliabilism—her emotional episode would not be regarded as a reliable cognitive process. I will not argue that emotions are necessary or sufficient for acquiring *all* kinds of moral understanding. However, I will argue that *emotional acquaintance* with tokens of morally appraised types of actions is indeed *necessary* for acquiring one type of moral understanding: the highest level. Hence, this is another respect in which Hills' view would not be compatible with my account of acquisition of moral understanding. I will describe this kind of emotional acquaintance in chapter V.

[22] Since the view of emotion that I will defend has a necessary feeling component, here I am understanding 'feeling' and 'emotion' as equivalent.

There will be another important distinction between Hills' view and mine. Hills also argues that in moral epistemology, moral understanding-why possesses higher value than moral knowledge-why mainly due to the fact that performing actions relying on our moral understanding warrants the *moral worth* of our moral actions, as well as the development of a virtuous character (Ibid: 106-119). However, the discussion of whether Mary's actions inside and outside the laboratory are morally worthy is out of the scope of this thesis.[23] In the next subsections, I will present some of Sliwa's objections to Hills' theory of moral understanding, and focus on applying Hills' theory of moral understanding to Mary's case.

ii.iii The Moral Reasoning Claim

In the previous subsection, I briefly introduced the *moral reasoning claim* as the claim that to possess moral understanding was to be capable of giving and following explanations why an action is either right or wrong (i.e., abilities i-vi above). The moral reasoning claim seems to liken the capacity to exercise these abilities with possessing the capacity for moral understanding-why (Sliwa: 541). Sliwa argues that there are three ways in which this is problematic:

(i) The moral reasoning claim confuses having moral understanding with having the ability to articulate it, but they can come apart. For example, Mary has both moral understanding and the capacity to articulate it through her moral reasoning abilities. On the other hand, Mary's sister Jane[24] who is a morally good person, who does the right thing because it strikes her as the right thing to do and responds to the morally relevant features of moral situations does not have the capacity to articulate it. Even if it would seem right to describe Jane as someone who definitely has moral understanding, Jane might not be very good at explaining her actions, nor at drawing conclusions about different moral scenarios.

(ii) The moral reasoning claim fails to allow the possibility of other kinds of moral understanding-why. For example, firm convictions of whether something is right or wrong such as 'helping those in need is right because I am sure it is right,' 'stealing is wrong because I am sure it is wrong.'

(iii) The moral reasoning claim downplays the significance of other cognitive processes in our acquisition of moral understanding-why. For example, first-hand emotional experiences with moral facts such as experiencing empathy when witnessing an act of bullying; after

[23] If pressed, my view would account for the moral worth of an agent's actions if they mirrored those of the virtuous person, who McDowell has described as having a "reliable sensitivity to a certain sort of requirement which situations impose on behaviour" (McDowell 1979: 142).

[24] Unlike Mary, Jane was not part of a moral experiment nor was she born in a moral laboratory, nor did she ever spend time in one.

> such an experience we might understand why bullying is wrong without being able to articulate a good explanation why. (Ibid: 541)

For these reasons, the moral reasoning claim appears to be a narrow conception of moral understanding. I argue that Hills' list of reasoning abilities is one way among many in which moral understanding can be acquired. To show this, I will now analyse Mary's case in terms of Hills' account of moral understanding.

It was expected that when Mary left the laboratory and she began to lead a normal life performing moral actions, she would make use of her moral understanding. Given her training and her reasoning abilities acquired inside the laboratory, Mary did possess the kind of moral understanding that Hills has described, even to a great extent, when she left the laboratory. I will now describe some of the ways in which both her capacity for moral understanding and the moral understanding she possessed improved due to her new emotional experiences, and some of the ways in which these experiences affected her negatively. The account of the following results will be brief and imprecise. I will discuss the nature of emotions and the way in which they can play a positive epistemic role in the acquisition of moral understanding in the rest of the chapters of this thesis (in particular in chapters III, IV). In chapter V, I will provide a detailed account of the difference between Mary's emotionless moral understanding, and her moral understanding after experiencing emotions.

Results that Showed Improvement in Moral Understanding After Experiencing Emotions

When Mary started her life outside the laboratory and she was no longer devoid of emotion, she began to experience the same kind of emotions that all of us experience in normal conditions (e.g., *sadness* when things seem unfortunate, *love* for family and friends, *anger*[25] when she perceives she is being wronged, *joy* when things seem to go well, etc.). But most importantly, she began to experience emotions that are often associated with moral judgements and moral actions. For example, *guilt* when she wronged someone else, *empathy* when she witnessed her friends suffering, *compassion* when she helped those in need, *indignation* when she observed acts of discrimination, etc. Instead of merely being a spectator and a student, some of her new emotional experiences *motivated* Mary to perform morally right actions.[26] Often, guilt motivated her to repair her wrongdoings, and empathy, compassion, and

[25] The case of anger is complicated. Many assume that it is a moral emotion, but others argue they can't tell the difference between moral anger and other instances of it (Maibom 2020). I will discuss what would make some emotions typically moral in chapter VI.

[26] It seems right to me that before leaving the laboratory, Mary mainly had knowledge-that. Or, she just had limited knowledge-how given that her only ability knowledge was restricted to moral reasoning. However, my position on the debate between intellectualists and anti-intellectualists about knowledge-that and knowledge-how is neutral.

indignation often motivated her to provide help. Hills' view would account for right motivation provided by emotional responses as a sign of virtue (2009: 108). Good motivation and good judgement are the two crucial components of virtue (Ibid: 108). And so, a courageous person (i.e., someone who experiences courage in the right way) is motivated to face danger whenever she grasps that it is worth doing so (2009: 109 footnote 28).

Mary realised that at least some of her emotional episodes led her to have new or deeper insights about why certain actions are right or wrong. Mary could now provide more and even better explanations about the supposed rightness and wrongness of different actions, and some things that she had learnt and understood from before made even more sense or were reinforced. For example, after lying to her new best friend Bob, she noticed that their friendship changed for the worse. She already knew that lying—among other reasons— is wrong due to the fact that the loss of trust within a community eventually leads to negative social behaviours that interfere with people's happiness. Mary added to her explanations that the guilt that she experienced after lying to Bob reinforced her understanding of the wrongness of lying, and one of her conclusions was that 'lying is also wrong given that it damages friendships.'[27] In other words, Mary began to make more use of her capacity of moral reasoning. Hills would say that in those cases where Mary is showing virtuous character, her emotional responses can lead her to can recognise moral reasons, just like an honest person grasps why honest actions are important and why she ought to perform them (2009:109 footnote 28).

Mary's improvement in her moral reasoning and her motivation to perform right actions at least in part due to her newly experienced emotional episodes, can also be accounted for on Hills' view on virtue. Hills argues that agents with moral understanding are able to perform morally worthy actions from a virtuous character (i.e., good motivations that respond to moral reasons).[28] She says:

> 'Being a good person is not just about what you do. Reliably acting rightly is a part of having a good character, of course, but a good, virtuous person is someone whose whole self—her thoughts, decisions, *feelings, and emotions* as well as her actions—is structured by her sensitivity to morality. This does not mean, of course, that a virtuous person is so consumed by morality that she thinks of nothing else. But it does mean that she is responsive to moral

(For anti-intellectualist positions see Ryle 1949; Carter and Pritchard 2015; for intellectualist positions see Stanley & Williamson 2001; Brogaard, 2008).

[27] Some authors such as Prinz (2007) and Vanello (2020) argue that emotional experiences are necessary to acquire evaluative concepts. I will not be arguing for that claim, since it is made clear by the experiment that I am discussing that Mary acquired the evaluative concepts of rightness and wrongness inside the laboratory, as well as many other thick concepts. However, in chapter V, I will explain—partially by following Goldie (2002)— in which ways Mary's understanding of morally evaluative concepts is enhanced or improved by her new emotional experiences.

[28] According to Hills, 'Good motivation is essential to having a good character, but it is obviously not sufficient. To be virtuous, you have to care about helping others and telling the truth, but you cannot be fully virtuous if you consistently mistake what is just, honest, or kind. Good judgment is a second crucial component of virtue.' (Ibid: 108) This relates to the kind of awareness involved in good judgement previously mentioned in section ii.ii.

considerations in all aspects of her character, whenever they are relevant.' (2009: 112, emphasis mine)

Hence, provided that they do not misfire, Mary's new experience of emotional episodes could allow her to become more virtuous and more able to complement her good judgement with good motivations.[29]

In sum, these results show how Mary's improvement in moral motivation and moral reasoning caused at least in part by her emotional experiences led to an overall improvement of her moral understanding. According Hills' view, it is possible to argue that her new emotional experiences have brought her even closer to developing a virtuous character. Hence, Hills would have to accept the claim that Mary's capacity for moral reasoning improved after her capacity for experiencing emotions was restored. However, if inside the laboratory Mary had *full understanding* given that she fulfilled moral reasoning abilities i-vi to the greatest extent, then Mary's improvement in moral understanding after experiencing emotions could not be explained in light of Hills' view unless it was in terms of virtue. It seems that by prioritising the capacity of moral reasoning, Hills is disregarding the significance that the capacity of experiencing emotions has in the acquisition of moral understanding.

Moreover, as I will later argue, Hills' list of reasoning abilities is incompatible with the possibility of gaining what I will define as *the highest level of moral understanding*. According to the view that I will introduce in section iii of this chapter, and fully elucidate in chapter V, the highest level of moral understanding would consist in part in a first-hand emotional experience with tokens of morally appraised types of actions. Hence, for Mary to *fully understand* the rightness of, say, helping war refugees, she would need to be adequately emotionally acquainted with the action of helping war refugees (e.g., to undergo an episode of compassion associated to the act of helping war refugees first-hand).

Results that Showed that Emotions Negatively Affected Moral Understanding

Mary also noticed that at times her emotional episodes clouded her judgements. For example, the first time Bob lied to her, Mary experienced so much anger that she yelled at him in a disproportionate and violent way. Her anger led her to judge that Bob was a morally bad person, and that he deserved a harsh punishment for having lied about his plans on Friday night. Hence, whenever her emotional episodes misfired, they motivated her to perform morally wrong or incorrect actions. Often, guilt motivated her

[29] Hills also acknowledges that in order to have some virtues, it is not necessary that one forms explicit moral beliefs. She says that, for example, it may be sufficient to be generous and judge that someone else is in need and that one could help them and go on to do so (2009: 109 footnote 28).

to mistake some of her actions as wrongdoings, and empathy, compassion, and indignation often motivated her to provide help either when it was not needed, or sometimes the help that she provided weakened the character of those she helped.

Due to other emotional outbursts, Mary also realised that she had to learn how to regulate her emotional responses. For instance, she would try to meditate and seek moral advice.

Also, Mary realised that she could not always explain all of her new moral insights to others, or at least not as clearly as she could explain things inside the laboratory. For example, after experiencing empathy towards Bob when he became a victim of bullying, she realised that although she could provide all sorts of explanations as to why bullying was wrong, she thought she still could not convey how wrong she thought and felt bullying was. It seemed that her strong episode of empathy rendered her unfit to provide more explanations. According to Hills' view, not being able to provide an explanation of the wrongness of lying would be an indication of a decrease or a negative effect on moral understanding.

There was yet another problem that Mary encountered. Mary realised that often, even though sometimes she was aware of the right actions that she was supposed to perform (e.g., share her meal with Bob when he did not have the money to buy any food), she could not bring herself to carry out the right action (e.g., in this case, the generous action of at least offering to share her meal with Bob). Mary noticed that whenever she failed to perform what she understood to be the morally right action, it was often because of her seeming failure to react emotionally in the appropriate way. In other words, at times Mary did not seem *to care* about performing the right thing (or care to avoid doing the wrong thing).

All of these and other problems can be easily diagnosed by Hill's account. As discussed above, emotions are required for virtue (2009:112), but they are not required for moral understanding (Ibid: 103 footnote 19). Hills points out that moral advice is extremely important in the context of deciding on moral questions, 'not just because they are inherently hard but also because your own desires, interests, and *emotions can bias you and lead you astray*' (2009: 123 emphasis mine). Roughly put, according to Hills' view, adequate emotional responses and good motivations only come in the picture of virtuous action but not as cognitive mechanisms that in themselves can provide moral understanding. Hence, given the ways in which Mary's new emotional episodes can lead her astray, it is easy to see why they are being excluded among Hills' six reasoning abilities required for moral understanding. I will argue in chapters V and VI, however, that in those cases where emotions do not misfire, they can increase both our capacity and level of moral understanding.

In the following section, I will describe Sliwa's Moral Knowledge Account, and briefly introduce a novel form of moral epistemic sentimentalism that attempts to highlight the importance of the role that emotions can play in the acquisition of moral understanding. The upshot of this thesis will be to show that Mary's moral emotional experiences improved her capacity and level of moral understanding

despite her having excellent reasoning abilities, and allowed her to fully understand the rightness and wrongness of certain actions.

iii. The Moral Knowledge Account and a New Moral Epistemic Sentimentalism

In this thesis, I take moral knowledge to be knowledge that employs moral concepts. For example, knowing that an action is right or wrong, knowing what the right or wrong thing to do is, knowing why an action is right or wrong, and so on (Sliwa 2017: 546). My aim in this section is to describe Sliwa's Moral Knowledge Account of moral understanding in order to compare it to the view of moral understanding that I will develop throughout subsequent chapters. In section ii.i I have already endorsed Sliwa's reductionism, according to which instances of understanding why *p* can be achieved depending both on whether agents know why *p* and on *how much* they know about why *p*:

> An agent understands why *p* if and only if she has a sufficient amount of knowledge why *p*. (Ibid: 530)

According to Sliwa, moral understanding is the ability to know right from wrong (Ibid: 523). Sliwa's moral knowledge account (MKA) aims to describe what the capacity of moral understanding consists in. The MKA suggests that the capacity of moral understanding is the ability to acquire moral knowledge (Ibid: 546). In other words, the ability to acquire moral knowledge is constitutive of moral understanding. Hence, an agent has moral understanding if and only if she has the ability to acquire moral knowledge (Ibid: 546). This ability should not be mistaken with mere capability,[30] in the sense that, for example, Mary having the capability of experiencing morally emotional episodes outside the laboratory will not warrant that she will undergo emotional episodes that will always lead her to acquire moral knowledge. If Mary's (or anyone's) capacity to react emotionally towards *prima facie* moral actions were very flawed, then she would not possess (or she would possess it to a very low extent) the ability to emotionally acquire moral knowledge.[31] Successfully exercising the capacity of moral understanding results in achieving instances of moral knowledge (Ibid: 547). Additionally, this ability to acquire moral knowledge works correctly when an agent is able to identify the *morally relevant features* of certain situations (or 'moral evidence,' as Sliwa calls it) (Ibid: 546-547).

In my view, the morally relevant features of a certain action, for example, would be the particular aspects which determine whether such action is either *prima facie* morally right or wrong. What morally

[30] Sliwa provides the example of what would entail her actually having the ability to speak French: 'I do not have an ability to, say, speak French merely because I could, in principle, learn to speak French, were I to enrol in a French class. To have the ability, I, in fact, need to have the relevant psychological mechanism that responds to French utterances.' (2017: 546 footnote 47)

[31] However, in chapter III I discuss some ways in which our emotional reactions can be regulated and therefore improved (see Elgin 2008).

relevant features consist in according to my view will be elucidated further throughout the rest of the chapters of this thesis.

Importantly, having the ability to engage in moral reasoning will be part of an agent's capacity of moral understanding, but it also allows for various others ways in which an agent can achieve instances of moral knowledge. In other words, agents may acquire moral knowledge by different cognitive mechanisms such as perception, imagination, intuition, affective responses, moral reasoning, etc. (Ibid: 547-548).

Below are three further central claims of the MKA that will also be compatible with my Moral Epistemic Sentimentalism (Ibid: 548):

(i) Although an agent might not able to adequately engage in moral reasoning, they might compensate by, for example, being particularly affectively attuned to morally relevant features of the situation.

(ii) Agents can have cognitive abilities to different degrees, and so what will ground the capacity of moral understanding will vary from agent to agent.

(iii) First-personal experiences are important to moral understanding. First-personal experiences provide a richer conception of the right or wrong-making features of certain actions (e.g., you learn why prisons can be dehumanizing to a higher degree by visiting a prison yourself).

Nonetheless, what will distinguish the *Moral Epistemic Sentimentalism* that I will develop from Sliwa's MKA is the emphasis on the importance of the role that emotions can play in the acquisition of moral understanding. Such emphasis is accounted for by (i) an elucidation of the components of emotion and how they can lead to an increase in moral understanding, (ii) a description of the way in which emotions direct attention, (iii) a view of emotional acquaintance that explains the possible acquisition of the highest degree of moral understanding of tokens of types of morally appraised actions, and (iv) a view of emotional episodes that can change our moral perspectives. This view will also provide a novel explanation of the reason why deferring to pure and impure moral testimony is problematic. The claims that ground my Moral Epistemic Sentimentalism (MES), and which I will develop in the following chapters are the following: (a) instances of moral understanding-why are reducible to instances of moral knowledge (b) moral understanding comes in degrees (c) emotions are sources of salience and can direct our attention to the morally relevant features of morally appraised actions, and (d) emotions can provide moral understanding in a distinctive way given their components.

Hills' account of moral reasoning abilities is similar to my MES, in the sense that both provide a detailed description of a way in which the capacity of moral understanding can be exercised. Also, a way to measure the degree of someone's understanding could be by assessing to what extent do they possess abilities i-vi. However, to measure an agent's degree of moral understanding would also involve, as

Sliwa's MKA suggests, taking into account how much knowledge an agent possesses. My Moral Epistemic Sentimentalism differs from Hills' view in that it is a view of moral understanding that endorses reductionism instead of non-reductionism. Also, it does not make the moral worth of actions depend on an agent's moral reasoning abilities. It explains that the highest level of moral understanding consists in part in a first-hand emotional experience with certain morally appraised actions instead of fulfilling moral reasoning abilities i-vi to the greatest extent. In chapters V and VI, my MES will explain why at least in some cases, Mary's level moral understanding improved after experiencing certain emotional episodes.

Since deferring to moral testimony is one of the main ways in which agents seem to acquire moral knowledge and moral understanding, in the next chapter I will discuss some ways in which deferring to pure and impure moral testimony can be problematic. My aim will be to argue that the reason why deferring to pure and impure moral testimony is problematic is *epistemic*. According to MES, deferring to pure and impure moral testimony does not lead to the acquisition of the highest level of moral understanding.

Conclusion

In this chapter I have introduced the hypothetical thought experiment of Moral Mary. The experiment is meant to investigate whether Mary's moral understanding is transformed after she leaves the moral laboratory and is able to experience emotions. I also endorsed Simple-K reliabilism as an analysis of knowledge, and argued for the possibility of moral knowledge.

I described Hills' account of moral understanding, and argued that it is not compatible with the claim that emotional acquaintance is required for the highest level of moral understanding. I showed that, whilst Sliwa's MKA of moral understanding does regard emotions and first-hand experiences as an essential part of moral understanding, it does not provide an explanation of the components of emotions and other ways in which emotions play a crucial role in the acquisition of moral understanding, which my Moral Epistemic Sentimentalism (MES) aims to do. I introduced MES as a view that emphasises the importance of the role that emotions can play in the acquisition of moral understanding, and I will continue to develop it in the rest of the chapters of this thesis.

Chapter II

Moral Testimony and Emotional Experiences

My aim in this chapter is to argue that forming moral judgements on the basis of moral testimony is problematic, mainly for the following reason: *the understanding that it provides is not of the highest level.*

In contrast to this view, Guy Fletcher has argued that forming moral judgements on the basis of moral testimony is problematic due to the fact that moral sentiments—which are intimately related to moral judgements—are at least difficult to form on the basis of pure, direct, testimony (Fletcher 2016). I will argue that Fletcher's view fails to adequately characterise the problem with moral deference. I will suggest instead that a key problem with forming moral judgements on the basis of testimony is *epistemic*. The agent who defers to the pure or impure moral testimony of somebody else does not acquire full understanding of the moral rightness or wrongness of the action in question. This is because the agent lacks a first-hand emotional experience with the action in question.

In the first section, I introduce the most widely accepted pessimistic theory which finds moral testimony problematic (*moral pessimism*), but mainly focus on describing Fletcher's version of this view. In the same section I also describe Laura Callahan's Affect and Motivation view (2018), which aims to provide novel reasons to hold moral pessimism. In the second section, I elucidate Daniel Wodak's optimistic argument (2019) for approving on the basis of moral and aesthetic testimony, in order to criticise Fletcher's pessimism, as well as moral pessimism in general. Thirdly, I briefly develop an account of moral emotional experiences, and claim that in contrast with extant pessimistic views, this account is an essential part of the best explanation of the epistemic problem with deferring to the moral testimony of others. Whilst I agree with Wodak's claim that deferring to the moral testimony of others can provide us with reasons to morally approve or disapprove, I challenge his optimism by arguing that having reasons to approve or disapprove is not sufficient for possessing the highest degree of moral understanding.

i. Why is Moral Deference Problematic?

Deferring to moral testimony seems to be one of the main ways in which moral knowledge is acquired. I take moral testimony to be testimony with explicitly moral content (Hills 2009: 94, footnote 1).

As Hills puts it, moral testimony might be testimony about what is morally right or wrong, for example, "It is morally right to perform acts of charity," or testimony that uses thick ethical terms (e.g., "It is generous to help those in need"), or about our moral reasons for action (e.g., "I have moral reason not to steal") (Ibid: 94). Although my view of the way in which deferring to the moral testimony of others is problematic includes the previous types of moral testimony, the discussion in this thesis mainly targets testimony about actions being either right or wrong. In this chapter, I will understand *moral deference* as the activity of forming and ceasing to hold true (or false) beliefs about supposed moral actions on the basis of the moral testimony of others. For example, children usually form their beliefs about the supposed rightness or wrongness of some actions on the basis of what their parents tell them (e.g., "It is wrong to lie"). In other words, children *defer* to their parents about the *prima facie* moral rightness or wrongness of certain actions. Inside the moral laboratory, Mary learned from the testimony of moral experts that, for example, taking lost wallets to the police station is 'right', and that stealing someone else's car is 'wrong.' In principle, forming a moral judgement based on someone else's testimony, usually someone whom we trust, respect, regard as an authority, etc., seems unproblematic (Fletcher 2016: 47). However, in this chapter, I will discuss some ways in which moral deference appears to be problematic.

There are two types of moral deference, *pure* and *impure* (McGrath 2009; Fletcher 2016). *Pure* testimony relates some supposed moral content without any ground. For example, Jane tells Mary that lying is wrong, and Mary defers to Jane. On the other hand, the content of *impure* moral testimony includes some morally relevant descriptive fact. For example, Jane tells Mary that lying is wrong given that it damages interpersonal relationships; Mary defers to Jane, and is able to provide Jane's explanation why lying is wrong.

In the case of pure testimony, there are at least two ways in which the receiver (in this case Mary) can *rely* on the testimony of the giver (in this case Jane). Once the giver shares putative moral information with the receiver, the receiver can rely upon this testimony either *directly*, or *indirectly* (Fletcher 2016:51):

Directly: X comes to hold the moral judgement that P solely on the basis of Y's testimony that P. For example, Mary can come to judge that 'lying is wrong' *solely* because Jane told her that lying is wrong.

Indirectly: X comes to hold the moral judgement that P partly on the basis of Y's testimony that P. For example, Mary comes to judge that 'lying is wrong' because Jane told her so, and because she already knows, say, that a maxim permitting lying would not pass Kant's universalizability test.

For now, I will only focus on problematic cases of pure and direct moral deference.

i.i Moral Pessimism

It is widely accepted that moral deference is problematic in the sense that it leads to a deficiency in moral understanding, which ultimately leads to actions that are not as morally praiseworthy as actions that are performed for the right reasons (Nickel 2001; Hills 2009; McGrath 2011, 2019). Hopkins has defined this claim as *unusability pessimism* (2007:614). However, for clarity purposes I will call this view *moral pessimism*. Moral pessimism belongs to the category of views according to which moral testimony may transfer moral knowledge,[32] but not moral understanding. Additionally, the actions that might result from this knowledge would not be as morally praiseworthy.

Hence, moral pessimism seems to involve at least two claims: (i) moral understanding is not acquired via pure, direct moral deference, and (ii) the moral worth of moral actions performed without moral understanding is undermined. Proponents of moral pessimism[33] such as Alison Hills, argue that moral understanding why-*p* (where *p* is a moral proposition) is not achieved through deference (2009: 97). Put simply, this kind of pessimist argues that at best, impure moral testimony can provide knowledge-that and also knowledge-why. Here I will only discuss Hills's pessimism about moral testimony briefly. Recall that according to Hills, moral understanding why *p* involves a set of at least six abilities to:

(i) follow an explanation of why *p* given by someone else; (ii) explain why *p* in your own words; (iii) draw the conclusion that *p* (or that probably *p*) from the information that *q;* (iv) draw the conclusion

[32] Another form of pessimism is one held by some non-cognitivists who deny the possibility of moral knowledge—and therefore the transmission of it—called *unavailability pessimism* (Hopkins 2007:614). However, I will not be concentrating in this kind of criticism, since I have argued in the previous chapter that the possibility of moral knowledge is not only accepted by realism but also by some anti-realist metaethical theories. After all, we frequently talk of moral knowledge, and we seek it in our daily lives. Nonetheless, other forms of *unavailability pessimism* might accept the possibility of moral knowledge, but deny that moral knowledge is transmissible by moral testimony. As we will see later in this chapter, some authors will argue that the impossibility of transmission of moral knowledge via testimony is psychological.

[33] Authors that defend moral pessimism can differ in their claims. For example, Hills (2009) and McGrath (2019) argue that moral knowledge can be transferred via moral testimony, but not moral understanding. However, these authors do not seem to hold the same account of moral understanding. Whilst Hills argues that moral understanding necessitates at least six abilities to show 'understanding why' a certain moral judgement is morally right or wrong, McGrath does not seem to argue for the same picture of moral understanding. Still, they both agree that forming moral beliefs based on the say-so of others does not 'put one in a position to do the right thing for the right reasons' (McGrath, 2019:93). Also, the moral competence of the receiver of the moral testimony in both cases seems to be measured by her accuracy of arriving at moral judgements *autonomously* (McGrath, 2019:156, Hills, 2009:101). It is important to note that the idea here is not that moral understanding is equivalent to having independent moral reasoning.

that *p*' (or that probably *p*') from the information that *q*' (where *p*' and *q'* are similar to but not identical to *p* and *q*); (v) given the information that *p*, give the right explanation, *q*; (vi) given the information that *p*', give the right explanation, *q*'. (2009: 102)

But these abilities are generally difficult to acquire, and they are certainly hard to acquire via pure, direct testimony (Ibid: 120). Moreover, on Hills's view all moral actions that are performed without moral understanding lack moral worth. For Hills, certain actions are morally worthy only if they are right actions performed for the right reasons (Ibid: 113). For example, suppose Mary defers to Jane's moral testimony that lying is wrong, and Mary ends up believing that lying is wrong solely on the basis that Jane told her so. Later on, Mary decides not to lie to her friend Bob. Clearly, Mary has done the right thing by not lying to her friend. However, when we ask Mary about her reasons for not lying to Bob, she simply replies 'because Jane said that it is wrong to lie.' Even if Jane is highly trustworthy, and a highly reliable moral agent, Hills would argue that Mary's decision not to lie to Bob is not morally worthy. Presumably, Mary's reason for avoiding lying in this case is not the right kind of reason. Mary would have avoided lying for the right reason, for example, if she had done it for the reason that it damages relationships. Also, morally worthy actions require that you are oriented properly 'not just in your outward actions but in your motivations, your choices, and your beliefs too' (Ibid: 117). Roughly put, according to Hills's moral pessimism, the set of cognitive abilities involved in moral understanding are hard to develop and so they are not generally acquired via pure moral deference. Furthermore, the actions that result from pure moral deference are usually not performed for the right reasons and so they lack moral worth or at least are morally defective in some way.

I will be discussing moral pessimism further throughout this chapter, but for now I will move on to Fletcher's view on what is problematic with moral deference in the next subsection. My general aim will be to argue that views such as Hills's and Fletcher's are not targeting the most significant problem with moral deference. Secondly, in section ii I will turn to Wodak's optimistic theory of reasons for approving and disapproving on the basis of normative testimony, in order to argue against both *moral pessimism* and Fletcher's *psychological pessimism*. Thirdly, I will introduce Callahan's version of moral pessimism to later compare it with my *epistemic pessimism*. Finally, despite agreeing with Wodak's objection to the previous kinds of pessimism, I will present a new problem with moral deference, as well as the reasons why Wodak's observation does not apply to it.

i.ii Fletcher's Psychological Pessimism

Pure and direct moral deference (henceforth just 'moral deference') seems more problematic than pure and direct deference in many other domains.

For example:

A1) 'Where's the stadium?'

B1) 'It's on 21st street. Ruth told me.'

A2) 'Is it wrong to watch boxing?'

B2) 'It's wrong to watch boxing. Ruth told me.' (Fletcher 2016:53)

B1 seems like a good reply, whereas according to Fletcher, the utterer of B2 would probably be interpreted as lacking sincerity because their utterance sounds like a joke (Ibid: 53).

In cases like the one above, Fletcher argues that the problem with moral deference consists in a *psychological difficulty* involved in the formation of moral judgements. This kind of case does not concern lack of moral understanding arising from deference, nor the undermining of moral worth of subsequent actions, as argued by moral pessimists (Ibid:52-53). Fletcher's pessimism would belong to the *affectively defective category*, given that it locates the problem with moral deference in an unlikely formation of sentiments. I will call Fletcher's view 'psychological pessimism.'

Before introducing Fletcher's explanation of this psychological difficulty, there is yet another important distinction which involves two instances in which pure and direct testimony can operate (see McGrath 2011:114):

Isolated: B is told by A that C performed a wrong action without learning any further details. B judges that C did something morally wrong, solely on the basis of A's testimony.

Change of Mind: B possesses further knowledge of the wrong action that C performed, and she doesn't question whether what C did was permissible or not, but B feels inclined to judge that C had certain reasons to act as she did, and thus did something morally permissible. However, after hearing A's testimony that C did something wrong, B decides to judge that C did something morally impermissible, solely on A's testimony.

For now, I will only focus on *isolated* cases of pure and direct moral testimony, but *change of mind* cases are also relevant in the explanation of Fletcher's view. I will address change of mind cases in section ii.i.

Fletcher's principle for explaining the psychological problem present in the case of B2, can work both for non-cognitivists who do not deny moral knowledge (and who think that moral judgements are sentiments), as well as for sentimental realists.

SENTIMENTS DEFERENCE DENIAL (SDD*): Moral sentiments[34] are *at least difficult* to form on the basis of pure and direct testimony (Fletcher 2016:60).

Consider how SDD* results in a non-cognitivist explanation of the problem with pure and direct moral deference:

a) *Isolated Testimony*: Bob tells Mary that John did something wrong, and Mary judges that John did something wrong.

Fletcher explains that unless Mary enriches the content of the testimony somehow, she will have insufficient information to form moral sentiments towards John or his action. Mary might be in a position to know that John has done something which she would feel anger, resentment, blame, or some other sentiment in response to, but she will not thereby form such an attitude. She will thus, given SDD*, be unable to form the moral sentiments constituting the moral judgement according to non-cognitivism. It then follows from Fletcher's view that if Mary says she has formed the sentiment that constitutes the moral judgement towards John's action (e.g., indignation), solely on the basis of Bob's testimony, it seems that Mary is being insincere, or is confused.

Fletcher also argues that one can use SDD* to explain the problem with pure moral deference on the supposition of a kind of realism: sentimental moral realism. Here is how he describes this realism:

> 'Suppose that moral judgements are wholly constituted by beliefs but that those beliefs are about the *fittingness* of moral sentiments. For example, to judge that X acted wrongly in φing is to believe that it is fitting, for example, to resent X for φing. Such a view can be a form of moral realism. It holds that moral judgements are beliefs (beliefs about fitting sentiments), that there are moral facts (facts about fitting sentiments), that such facts do not themselves depend upon agents' sentiments, and it is compatible with the idea that we have moral knowledge. Call such a view "realist-sentimentalism" (2016:64).

On a view like this, there will be an internal rational pressure for one's sentiments to match up with one's moral beliefs (whether or not one's beliefs are true). In other words, it would be irrational to think it fitting to resent X for φing, and yet not resent X for φing. An additional pressure that seems to be at play here, is that according to the realist-sentimentalist the moral judgements that correspond to certain moral facts are usually associated with certain moral sentiments —and so in an external sense these

[34] As Tappolet notes, moral sentiments are often taken to be *'dispositions* to undergo a range of emotions, such as *guilt* if one has committed something deemed morally dubious, and *indignation* if the deed is someone else's' (2016: 5, emphasis mine). Hence, sentiments appear to be dispositional states (i.e., enduring) rather than *occurrent*, which is usually what distinguishes them from emotional episodes (Ben-Ze'ev 2010: 55). In this thesis I will be focused on emotional episodes, although I will use Fletcher's term 'moral sentiment' when discussing his view. My view of emotions departs from Ben-Ze'ev's in the sense that, I do not mark a special distinction between long-lasting emotions and sentiments. In my view, the terms 'emotion' and 'sentiment' can be interchangeable. Here I follow Cohon's interpretation that Humean moral sentiments are emotions in the present-day sense of that term (Cohon, 2008: section 7).

sentiments should be 'fitting'—. It is possible to use the same example as before, to see how SDD* can operate under the supposition of a realist-sentimentalist view:

a) *Isolated Testimony*: Bob tells Mary that John did something wrong, and Mary believes that John did something wrong.

Unless Mary enriches the content of the testimony, she will have insufficient information to form the appropriate sentiments towards John or his action, given SDD*. According to realist-sentimentalism, this can be explained due to the fact that moral sentiments are responses that we have to the properties of the relevant action or agent, etc. Without proper acquaintance with them, the attitudes are very unlikely to arise even if we know that they would, were we so acquainted. In cases of such deference, the problem stems from the receiver coming to form the judgement that negative moral sentiments are fitting whilst—according to SDD*— probably lacking such attitudes. Mary can judge that it is appropriate for her to feel negative moral sentiments towards John (e.g., indignation and resentment) without actually feeling them. The issue is not necessarily that Mary (i.e., the receiver) will be unable to form the *belief* that John did something wrong. Rather, the problem is that Mary will lack the desire-like moral sentiment mentioned as fitting in the content of the belief, given her lack of acquaintance with John and his action, and in light of SDD*.

In conclusion, Fletcher has identified pure and direct moral deference as being problematic at a psychological level. According to his criticism, there is a seeming oddness/implausibility involved in cases of moral deference, where the receiver claims to have formed the moral sentiments associated to the moral judgement of the giver. According to SDD*, due to the difficulty involved in the formation of moral sentiments present both in non-cognitivist as well as realist-sentimentalist[35] pictures of moral judgements, the receiver appears as confused, conflicted, or insincere rendering moral deference problematic. In contrast to what Fletcher claims, I will argue that even if in some cases there is a problem in the formation of moral sentiments, this will still not be the relevant problem with moral deference.

In the next section, I plan to introduce Daniel Wodak's optimistic account of approving on the basis of moral testimony in order to raise an important objection for both moral pessimism and psychological pessimism. In section ii.i, I will discuss two additional problems with Fletcher's view.

ii. Problems with Moral Pessimism and Psychological Pessimism

Daniel Wodak has argued that normative testimony (i.e., aesthetic and moral), gives us reasons for attitudes like approval (Wodak 2019:29). In this discussion I will put the case of aesthetic testimony

[35] Fletcher also mentions that SDD* applies to "hybrid" views, which hold that moral judgements are necessarily accompanied/partially constituted by moral sentiments or desire-like states (2016:63-64).

aside, since the kinds of pessimism that I have described so far only involve moral deference. Wodak questions the comparison between cases of pure deference and acquaintance, since it is often argued that only acquaintance provides us with strong reasons to approve of moral actions (Whiting 2015; Lord 2016).

Wodak argues that through the right kind of comparison cases, it is possible to find 'better' and 'stronger' reasons that would justify our approval or disapproval on the basis of pure deference. By the 'right comparison cases,' Wodak suggests comparing changes in the content of the *same kind* of testimony. He calls these cases 'minimal contrastive pairs' (Wodak 2019:6). Here is his example of how in a case of pure moral deference, it is plausible to find better reasons to form a moral judgement on the basis of a reliable person's testimony:

'Wei's case: Wei tells you that he saw a man do something wrong. Knowing Wei to be highly reliable, you disapprove of the man's act.

Zhao's case: Zhao tells you that he saw a man do something evil, monstrous, utterly reprehensible, and so on. Knowing Zhao to be extremely reliable, you disapprove of the man's act.

There are two key differences between Wei's and Zhao's testimony. The first one is that Wei's testimony provides less information on the description of what the man has done, and the second one is that Wei is *highly reliable* whereas Zhao is *extremely reliable*.

Wodak argues that if we compare Zhao's case with Wei's, there seem to be strong reasons to disapprove on the basis of Zhao's word (Ibid:12). He argues that it is even plausible to have strong reasons to disapprove of what that man did solely by deferring to Zhao. Wodak says that when we compare Zhao's case to ordinary cases of moral acquaintance, it is rarely the case that when witnessing a wrongful deed one carefully studies and analyses its features:

> 'Often, we disapprove of others' misdeeds on a much flimsier basis than that. And often we're justified in doing so. (One who denies this, risks being committed to an over intellectualised view of moral epistemology). If our disapproval of wrongful actions on the basis of ordinary acquaintance can be justified, surely the same can be said when we disapprove on the basis of extremely reliable testimony that actions are evil, reprehensible.' (Ibid:12)

Wodak uses the same argumentative strategy against moral pessimism. He characterises the moral pessimist view in this way (Ibid:23):

- NO MORAL REASON: [Pure] Testimony that φing is morally right cannot give you any moral reason to φ (Hills 2009; McGrath 2011).
- NO PRAISE: [Pure] Testimony that φing is morally right cannot give you sufficient moral reason to make it praiseworthy for you to φ (Hills 2009; McGrath 2011).

Here is the example that Wodak gives which shows that, by changing the reliability and the content of the same kind of testimony it is possible to argue not just for having reasons to approve on the basis of pure, direct deference, but also that it is possible to argue for the moral worth of such actions performed from it (Ibid: 25-26):

> *Pablo's case*: A spy infiltrates a villain's lab. Her highly reliable informant, Pablo, left her a message: "There's a big red button in the next room. It'll be hard for you to push it, but morally you must do so."
>
> > Say that against all odds, the spy manages to push the big red button. She saves the day. I already find it intuitive that she had at least some moral reason to do so; indeed, I find it intuitive that she is morally praiseworthy, even when we stipulate that she was acting solely on the basis of Pablo's testimony (supposing, as we have throughout, that the testimony was true). But we can bolster these intuitions by considering variants on the case:
>
> *Quentin's case*: A spy infiltrates a villain's lab. Her highly reliable informant, Quentin, left her a message: "There's a big red button in the next room. It'll be hard for you to push it."
>
> *Rae's case*: A spy infiltrates a villain's lab. Her fairly reliable informant, Rae, left her a message: "There's a big red button in the next room. It'll be hard for you push it, but morally you must do so."
>
> *Sami's case*: A spy infiltrates a villain's lab. Her extremely reliable informant, Sami, left her a message: "There's a big red button in the next room. It'll be hard for you push it, but morally you must do so."

In all of these cases, the spy pushes the button and saves the day, which was the right thing to do. When the moral content of the testimony is absent (as in *Quentin's case*), the spy seems to have less moral reason to push the button, as well as when the informant is less reliable (as in *Rae's case*). On the other hand, it seems that the spy had more moral reason to push the button when the informant was more reliable (as in *Sami's case*). Therefore, according to Wodak, moral testimony can give us at least some moral reasons to act, and sometimes even strong moral reasons. In the same manner, moral testimony can give us moral reasons which merit praise.

Similarly, Wodak argues against Fletcher, that if moral testimony gives us reasons for certain attitudes, it shouldn't be fishy or odd to form those attitudes in response to normative testimony. Recall Fletcher's psychological pessimism (SDD*): Moral sentiments are at *least difficult* to form on the basis of pure and direct testimony.

According to Fletcher, approving or disapproving solely on the basis of pure and direct moral testimony is problematic, given that the receiver might not come to form the judgment that certain moral

sentiments are fitting (in the realist-sentimentalist case), nor the moral sentiment that constitutes the judgement (in the non-cognitivist case).

Wodak aims to show in five steps that Fletcher's account is incomplete (Ibid:28). I have provided content to these steps, as an example of his criticism to Fletcher's view in its realist form:

(1) It is odd for Mary to believe that John did something wrong in response to Bob's testimony (given that she's not acquainted with other details of John's action).

(2) But (according to Wodak), Bob's testimony that John did something wrong provides good epistemic reason for Mary to believe John did something wrong.

(3) It is odd for Mary to believe that John did something wrong without (e.g.) disapproving of an X property of John's action (given that Mary is expected to experience an emotional response towards John's action).

(4) It is odd for Mary to (e.g.) disapprove of X in response to Bob's testimony that John did something wrong.

Wodak points out that there is a tension between (1) and (2) explained by (3) and (4); it is odd for Mary to both believe and disapprove on the basis of Bob's testimony that John did something wrong. However, if Wodak is right as implied by (2), we should also accept (2019:29):

(5) Bob's testimony that John did something wrong provides good reason for Mary to (e.g.) disapprove of X.

But Wodak says that, if we accept this claim the tension between (1) and (2) re-emerges now between (4) and (5) 'But if we accept this claim, the same tension we saw between (1) and (2) re-emerges, only this time between (4) and (5)' (Ibid: 29). Wodak claims that Fletcher's view does not provide an answer as to why it is odd or fishy (in this case) for Mary to respond to reasons to believe that John did something wrong on the basis of Bob's testimony, nor why is it odd or fishy for Mary to respond to reasons to disapprove of x by disapproving of x.

Wodak's optimistic view has shown that pure and direct moral deference can supply us with reasons to believe, act, and form affective attitudes of approval and disapproval. Through the right comparison cases, moral deference can even give us sufficient moral reason to make it praiseworthy to act appropriately in light of that testimony. Wodak's account provides us yet with another reason to reject moral and psychological pessimism.

In the following subsection, I will discuss two further objections to Fletcher's psychological pessimism.

ii.i Two Additional Problems with Fletcher's Psychological Pessimism

Fletcher describes two cases that intend to demonstrate the difficulty involved in the formation of moral sentiments that are associated with or constitute moral judgements. However, I argue that at least in the following two cases this difficulty dissipates or it is undermined due to two main reasons: (i) the receiver might transform their moral sentiments easily if they already possess the same emotional disposition as the giver, and (ii) the receiver might transform their moral sentiments easily if they deeply trust the giver. I argue that if receivers can easily transform their moral sentiments at least in those two cases, this shows that the difficulty in the formation of sentiments cannot fully account for the way in which moral deference is problematic.

Recall SDD*:

SENTIMENTS DEFERENCE DENIAL (SDD*): Moral sentiments are *at least difficult* to form on the basis of pure and direct testimony (2016:60).

In an example of *Change of Mind*, Fletcher presents a case where someone changes their mind from judging that capital punishment is permissible, to judging that it is impermissible. In a case like this, Fletcher considers that it is quite hard to imagine the receiver undergoing such a radical change in their moral sentiments, only because she was compelled to believe what the giver told her. However, it seems *at least possible* that receivers already have the relevant moral sentiments that can allow them to defer without much problem. For example, *S* might judge *X* wrong and disapprove of it. Nonetheless, *S* also slightly approves of *X* too. In other words, *S* has conflicting sentiments towards *X*. For example, perhaps *S* approves of *X* insofar as *X* is G, but disapproves of *X* insofar as *X* is *F*. Given this conflict, it is possible, in response to testimony, for the sentiment of approval towards *X* to overcome the sentiment of disapproval towards it. It is possible to see *S* coming to fully approve of *X* and judge it to be right, if they already had a slightly formed sentiment of approval towards *X*, and if they really trust giver *Y* who approves of *X*.

So, for instance, using Fletcher's example, it might be that *S* initially judges capital punishment as permissible and feel a sentiment of approval towards it (e.g., righteousness), but *S* might also at times feel a sentiment of disapproval towards capital punishment (e.g., indignation). Then *Y*, whom *S* really trusts, tells *S* that capital punishment is impermissible. It is then possible to imagine *S*'s feeling of indignation amplifying and stopping them from feeling righteousness towards capital punishment. Hence, after deferring to *Y, S* might end up judging capital punishment as impermissible. Moreover, if *S* has already felt indignant towards scenarios which involve unfair law enforcements (e.g., some extradition cases), it wouldn't be as difficult or odd for them feel different about the permissibility of capital punishment and change their mind.

In both cases of *Isolated* and *Change of Mind*, Fletcher seems to overlook that even if it seems at least difficult to form adequate moral sentiments via deference, such complication can disappear entirely if the receiver of the testimony already possesses the relevant moral sentiments associated with certain moral judgements, such that she has no moral sentiment left to form, and in this way, the door for gaining moral knowledge through pure, direct, testimony is open. So, also, if receivers already possess the moral sentiments required to hold certain moral judgements, or happen to have identical moral sentiments to those of the givers, Fletcher's theory of the way in which pure and direct moral deference is problematic no longer works.

Another worry with Fletcher's view, is that it becomes less intuitive once we are closely acquainted with the person we are deferring to. Even if it holds that pure and direct testimony may present a difficulty in the formation of moral sentiments, and even if it seems objectionable to defer directly on moral matters, the psychological oddness that Fletcher describes would appear to diminish once the person one defers to is someone whom we deeply trust. As Wodak has pointed out, it seems perfectly reasonable and psychologically possible for, for example, Mary, to form a moral sentiment by accepting Jane's testimony (since she has reasons to), but her personal relationship and affective attachment to Jane might also make the formation of such sentiments a lot easier.[36]

Additionally, another way in which it would not be as odd for Mary to form a moral sentiment by deferring to Jane would be explained by the phenomenon of *emotional contagion.* Emotional contagion takes place when people start experiencing similar emotions merely as a result of the association with other people (Stueber 2019: 4). For example, one can start feeling sad because other people around are sad, or one can start feeling anxious if surrounded by a crowd that is experiencing anxiety. Hence, emotional contagion also explains why when she defers to Jane, Mary can form the same moral sentiments as Jane (i.e., by merely being associated with her).

Hence, the previous examples show that Fletcher's explanation of what is problematic with moral deference is incomplete. It is possible to question the alleged difficulty in the formation of moral sentiments in those cases where the receiver has conflicting sentiments, or when they deeply trust the giver, or when the receiver shares the giver's general emotional disposition.

In the next subsection, I will introduce Callahan's Affect and Motivation view in order to highlight the relevance of affect and motivation to moral understanding (Callahan 2018). Although Callahan's view is another version of moral pessimism, it introduces a problem with moral deference that will be similar to the problem that I will elucidate. However, I will argue that Callahan's account of the way in which moral deference is problematic is also incomplete.

[36] Yet another way in which it would not be so odd for Mary to form the relevant moral sentiment when she changes her mind, would be if her general emotional disposition were very similar to Jane's.

ii.ii Callahan's Affect and Motivation View

Callahan offers different reasons for endorsing moral pessimism, or at least the moral pessimism of the kind that Hills endorses. Like Hills', Callahan's view of the problem with moral deference applies to cases of both pure and impure moral testimony, (unlike Fletcher, who only discusses cases of testimony with pure moral content). Recall that according to moral pessimism, moral deference is problematic given the lack of moral understanding that the receiver acquires from deferring to the giver. As a result of this lack of moral understanding, the receiver ends up performing actions with diminished moral worth. On Hills's account, having a thorough understanding of moral propositions (e.g., understanding why taking a lost wallet to the police station is right) is so important, given that it is considered to be necessary for reliable right action and for morally worthy action.[37] On this account, moral actions performed on the basis of deferred moral knowledge-why lack moral worth because such moral beliefs are not likely to be grounded on autonomous reasoning. However, Callahan argues that Hills's conception of moral understanding-why (henceforth just 'moral understanding') as a set of cognitive abilities seems to be incomplete, and therefore moral pessimism so far described by Hills does not accurately explain what is really problematic about moral deference (Callahan 2018:447).

Callahan's explanation for endorsing moral pessimism consists in two parts: (i) reconceiving understanding in the moral domain as a *richer* state that also comprises affective and motivational engagement with reasons as well as cognitive ability, and (ii) offering another account of the tension between understanding and deference (Ibid: 438). I will proceed now to describe in some detail (i) and briefly describe (ii). In section iii, I will explain why even if my conception of moral understanding is very similar to Callahan's, our project of incorporating the affective component in moral understanding differs in at least two important ways.

Callahan begins to argue for (i), noting that some metaethical views require that all moral beliefs are accompanied by emotional or motivational states (e.g., a kind of realist sentimentalism). However, she points out that it is also possible to adopt a weaker view whereby 'desirable or normative moral beliefs will be accompanied by appropriate affect and motivation' (Ibid: 447). Also, as mentioned before, on

[37] Despite granting that understanding moral propositions seems to be distinctively valuable in the sense that it appears to be necessary for reliable right action and/or for morally worthy action, Callahan points out that this should not be taken as the feature that distinguishes moral propositions from other propositions that do not seem problematic to take on testimony (e.g., Mary enjoys Bob's company) (Ibid: 444). This seems to be the case, since understanding certain non-moral propositions can be necessary for both morally right action and morally worthy action. For example, by understanding why Mary enjoys Bob's company—Bob is a loyal friend— Jane advises Mary to apologise to Bob for having lied to him, and we can say that Jane's action was right and morally worthy. Hence, Callahan argues, if it is correct that understanding propositions with non-moral content can be necessary for moral understanding, it can be argued that understanding moral propositions is not especially valuable only because of its being necessary for those types of action (Ibid: 444-445). Nonetheless, as mentioned in chapter I, I do not plan to engage in a discussion (or take a stand) regarding the conditions that would make a moral action worthy or unworthy.

some forms of non-cognitivism, mental states such as sentiments can constitute moral judgements. If this phenomenon can describe in some ways how we think about our moral judgements, proponents of moral pessimism should agree that at least part of what makes moral testimony problematic lies in the difficulty of transmitting appropriate affect and motivation (Ibid: 454). As seen in section i.i, Fletcher has provided an account to explain this difficulty. Further, others such as Enoch (2014) and Howell (2014) have also highlighted both the difficulty in the formation of appropriate affect via deference, and the general importance of moral sentiments in the formation of moral judgements.

Put simply, Enoch has argued that moral deference seems problematic because of a lack of emotional responses (on the receiver's part) to the morally relevant features of a situation (2014:254). In a similar vein, Howell argues that moral testimony is not generally a good way of acquiring appropriate emotional or motivational dispositions, but he also claims that an agent's moral judgements based on moral testimony are in tension with the agent's development of a virtuous character (2014: 402).

Hence, it would seem that according to these three authors (i.e., Fletcher, Enoch, Howell), a case of pure deference where—all things considered— an appropriate moral sentiment is acquired through moral testimony would not be regarded as problematic. However, Callahan argues that such a case seems paradigmatically problematic or 'fishy.' For example:

> 'FRED treats Daphne as his moral guru. He not only believes automatically that whatever Daphne says about morality is right but—because of his passionate admiration for and attraction to Daphne, his desire to be like her and to be approved by her—he is affectively and motivationally moved by the moral facts Daphne communicates to him.' (Callahan 2018: 448)

Besides pointing out that having a moral guru should be regarded in itself as odd or fishy, Callahan says that Fred's emotional and motivational engagement with Daphne's true moral testimony seems even more worrying than the alternative case where Fred would not form adequate moral sentiments by deferring to Daphne (Ibid:448). Enoch might respond that Fred's accurate emotional engagement is less of an achievement than Daphne's, who has responded to morally relevant features of situations *directly*. However, Callahan argues that it is not clear why we should regard Daphne's achievement as necessarily better only because we value correct or appropriate emotional responses towards moral facts (Ibid: 448-449).

According to Callahan, Fletcher would also have to explain why at least in some cases where we correctly acquire sentiments via moral deference, this might still seem odd or fishy. And so, says Callahan, the difficulty in the transmission/formation/acquisition of affective and/or motivational states through moral testimony does not seem to explain why moral deference is problematic (Ibid: 449).

Callahan reinforces this claim also by noting that it is possible to have cases of moral deference where appropriate emotional and motivational engagement paired with accurate moral understanding can seem

problematic. For example, when she defers to Jane, Mary understands why she must apologise to Bob for having lied to him, at the same time that she feels guilty and is moved to ask Bob to forgive her. Nonetheless, her guilt and her motivation may be mainly caused by her need to be constantly praised by Bob.[38]

Nevertheless, Callahan thinks that there seems to be an additional norm that requires not only that appropriate transmission/formation/acquisition of affective and motivational states together with moral understanding-why take place, but also that these states are unified (Ibid). Following Howell, Callahan argues that this lack of unification between adequate emotional, motivational, and cognitive states (i.e., understanding-why) in moral deference is explained by the lack of development of virtuous character. However, Callahan argues that there is a more specific way to explain what is problematic about moral deference by re-conceiving understanding as 'a richer state,' which I will now explain (Ibid: 450).

So far, as we have seen, the standard conception of moral understanding held by moral pessimism mainly involves the ability of grasping reasons (i.e., Hills's view). Callahan's proposal is to change this conception of moral understanding, and think of moral understanding as not only requiring cognitive abilities involved in grasping moral reasons, but also as an *affective* and *motivational* engagement with moral reasons (Ibid: 450). As I will explain in the following section, my view will also emphasise the relevance of the emotional component and motivational component in moral understanding. However, I will not be arguing for the significance of the emotional and motivational engagement with moral reasons in the same way as Callahan does. Callahan argues that these aspects are *morally* significant (i.e., important for moral worth and our conception of moral understanding), whereas I will argue in chapters III-VI that they are *epistemically* significant (i.e., important for the acquisition of moral knowledge and moral understanding). Callahan says:

> On my account, having understanding of 'p' still entails believing other propositions related to 'p', as well as seeing and being able to draw the connections between them. *But it also entails having the disposition to host emotions and motivations appropriate to the truth of 'p', as well as instantiating some broader, parallel affective and motivational dispositions*. For example, where 'p' is a particular moral judgement, having understanding of 'p' will entail having the disposition to feel similar emotions or motivations faced with a range of distinct, saliently similar cases (in which 'q,' 'r,' etc. become relevant). (Ibid: 450-451) (emphasis mine)

[38] Callahan provides the examples of Marjorie and Rehan, who are 'exemplar moral calculators' (i.e., their capacity for moral understanding is really good), but that nonetheless perform actions that are not truly morally worthy. Despite engaging emotionally and motivationally with their moral beliefs, they fail to be recognised as virtuous people given their deficiency in appreciating and responding to moral reasons. In this sense, Marjorie and Rehan would be just like Moral Mary recently after she has left the laboratory. Mary has acquired an excellent capacity for moral reasoning, but when she is no longer emotionally sedated, sometimes he still fails to respond adequately to the right moral reasons.

For Callahan, having understanding of 'p' involves a more holistic conception of the standard moral understanding-why 'p', where 'believing that p for the reasons that make "p" true' also requires appropriate emotions and appropriate motivations (Ibid:451). So, for example, in Callahan's re-conceived conception of moral understanding an affective and motivational engagement with moral reasons would look like this (Ibid: 451-452):

(i*) (target element) Mary believes that (*p*) (e.g., lying to her friends is wrong) and has the fitting affective/motivational response given that (*p*) (e.g., the disposition to experience guilt/make amends);

(ii*) (cluster element) Mary believes some relevant cluster of other propositions that *'p'* supports and/or are supported by '*p*' (e.g., lying is generally wrong because it involves betrayal, lying to members of your family is wrong), and also has fitting affective/motivational responses given those other propositions (e.g., the disposition to experience shame/apologise sincerely);

(iii*) (appreciative element) Mary cognitively appreciates the support relationships mentioned in (ii*) and her exhibiting the affective/motivational responses mentioned there is sensitive to those relationships; and

(iv*) (ability element) Mary has some requisite level of cognitive, emotional, and practical facility with '*p*'-related reasoning —e.g., Mary can not only see or appreciate connections but, e.g., perform inferences and thus be moved to action, or feel the import of explanations.

Callahan claims that with this modified conception of moral understanding, moral pessimism is better characterised. Whilst Callahan agrees that understanding moral propositions plays a particularly distinctive role in reliable and morally worthy action, and that it partially constitutes moral virtue, she worries that the conception of moral understanding as a set of cognitive abilities is 'too thin' to be regarded as an important component of moral worth and rightful action. She argues that according to this narrow conception of moral understanding, it seems that moral understanding is only valuable because it is a necessary condition on morally worthy action or virtue (Ibid: 452). However, once a more comprehensive picture of moral understanding is provided, according to Callahan, it is easier to argue for its moral value.

Finally, the second part of Callahan's explanation for endorsing moral pessimism consists in offering a new account of why acquiring moral understanding via moral deference is defective. Callahan argues that acquiring moral understanding via moral deference 'discourages' *further acquisition* of moral understanding (Ibid: 454). Callahan describes this deficiency in acquisition via moral deference as giving less reason to engage in further moral action and reflection. For example, after deferring to Jane, Mary adopts a settled view on the wrongness of lying and so—all things considered— she now has less reason to think about other reasons why lying is wrong. In other words, one of the results of Mary's

deference would be stopping her from acquiring greater moral understanding of the moral propositions that account for the wrongness of lying.

Callahan's example of how moral deference can discourage or disincentivize acquisition of moral understanding involves three different characters: Igor, Constance, and Bella (Ibid: 455). The three of them lack understanding of a true moral proposition *p*. Although Bella is in the best epistemic position to understand *p* in comparison with Igor's and Constance's, the processes whereby the three of them would have to acquire understanding come at a cost (Ibid:455). The efforts, time, and attention that it takes to study, inquire, and think about moral matters makes it difficult for different kinds of agents (with low, medium, and high levels of epistemic abilities)[39] to seek further moral understanding of *p* after accepting a (say) moral expert's testimony. Even if it were easier for Bella than for Igor and Constance to inquire more about *p* after deferring to the moral expert's testimony, she might have less reason to do so given her already high epistemic abilities. Therefore, according to Callahan, accepting a moral expert's testimony of a true moral proposition *p* would potentially discourage Igor, Constance, and Bella from acquiring moral understanding.

Hence, learning moral propositions via moral deference might disincentivize acquiring moral understanding, which according to Callahan further supports moral pessimism.[40] On this view, it can be argued that Mary's acquisition of moral understanding outside the laboratory when she defers either to Bob or to Jane is discouraged. This result is not particularly relevant for my discussion, although Callahan's novel explanation of defective acquisition of moral understanding will be important for my view in a way that I will develop in the next section. So far it is worth mentioning briefly that Callahan's attempt of broadening the moral reasoning conception of moral understanding is mainly helpful—on my view— to emphasise the relevance of emotion in our general view of what makes moral understanding and moral action valuable, and to introduce the role that emotions seem to play in our acquisition of moral understanding. However, I will argue that moral pessimism does not provide an account of the most significant epistemic problem found in cases of moral reference.

[39] Psychological differences are also taken into account in Callahan's example (e.g., Constance might be more curious to keep on finding out the reasons why *p* is right or wrong).

[40] It is also worth noting that Callahan's moral pessimism does not escape Wodak's criticism in the cases of pure and direct moral deference. Through the right comparison cases, the receiver can find some or strong reasons to defer to the giver, even if the understanding in question is 'enriched' by an affective component. Wodak can argue for the possibility of finding reasons for the formation of emotions and motivations just like he did against Fletcher's pessimism. However, Callahan might reply that finding reasons for the formation of motivations might involve a more complicated process.

iii. A New Epistemic Pessimism

Whilst I agree with Callahan's claim that leaving out the emotional and motivational component in moral reasoning results in a narrow conception of moral understanding, I will not be arguing for the same picture of moral understanding that she endorses. Roughly put, Callahan's conception of moral understanding consists in adding emotional and motivational components to Hills's reasoning abilities. I consider my conception of moral understanding to be even more comprehensive than Callahan's, since—as I argued in the previous chapter—moral understanding can be acquired in as many ways as moral knowledge can be acquired.

According to my Moral Epistemic Sentimentalism (MES), although moral knowledge can be acquired through emotional and motivational responses, it can also be acquired without these responses and still be regarded as equally valuable *qua* moral understanding (e.g., via moral reasoning or via epiphanies). Recall that MES is an account of moral understanding that emphasises the importance of the role that emotions can play in the acquisition of moral understanding. According to MES, moral understanding is graded. Hence, my conception of moral understanding differs from Callahan's, in that it provides an account of the highest level of moral understanding. Simply put, Callahan only argues that an agent with understanding that includes emotional and motivational elements will have the 'permanent disposition of character' required for virtuous action (2018: 453).[41] In contrast, MES explains that the highest level of moral understanding consists in part in an adequate *first-hand emotional experience with tokens of morally appraised types of actions*, and the lack of these first-hand moral emotional experiences in moral deference explains why moral deference is problematic. I will call this kind of emotional experience *first-hand morally emotional experience.* I define 'first-hand emotional experience' as the direct epistemic access to the relevant features of a given situation through the experience of an emotional episode. I will describe this 'emotional acquaintance' in chapter V. The kind of emotional acquaintance that I will be discussing in this thesis is the one that provides access to the morally relevant features of certain actions.

Callahan argues that moral deference is problematic given that it discourages further acquisition of moral understanding. However, my epistemic pessimism is more radical. I do not argue that moral deference is problematic because it can foster an attitude of complacency which can discourage the receiver to keep on acquiring moral understanding. Rather, I argue that moral deference can never, by itself, lead to the acquisition of the highest level of moral understanding. Below is a definition of my epistemic pessimism about moral deference.

[41] Callahan follows Aristotle on this conception of virtuous action. She says: "Aristotle held (2000: 1105a28–33) that for actions to be done virtuously they must be chosen deliberately, in the knowledge that they are right, and they must also 'spring from a fixed and permanent disposition of character'." (2018: 453)

Moral Epistemic Pessimism: Moral deference cannot provide the highest level of moral understanding, given the receiver's lack of a first-hand emotional experience with the morally relevant features of a certain action.

Those features in virtue of which certain actions are *prima facie* right or wrong would be the morally relevant features of the actions in question (e.g., violent comments and violent gestures would be the morally relevant features of the way in which someone behaves towards their partner, and which make their action *prima facie* wrong). In contrast with Fletcher's view, MES does not hold that moral deference is problematic due to the fact that the moral sentiments that are associated with the testimony, or that consist in the moral judgement of the giver cannot be formed by or transmitted to the receiver. Rather, the problem is that the receiver will lack the first-hand *phenomenological experience* that the giver has had. Generally, the receiver of a moral testimony would be someone who has not acquired the highest level of moral understanding-why of the act-token that the testimony is describing. A receiver who has already undergone the same first-hand emotional experience as the giver then would not be really 'deferring.' In other words, an agent who partakes in the testimony of an exact emotional experience that they have lived through, would not be —in a strict sense—a receiver.

Although, for example, Jane could describe to Mary in great detail what it was like—*emotionally*— to help war refugees, Mary would still not be able to understand the rightness of the same act-token of helping war refugees in the same way as Jane.[42] Jane could explain to Mary what it was like to experience *compassion* throughout the act of helping war refugees.[43] Jane could also describe in a very detailed manner different features of the act of helping war refugees that can help Mary understand further the rightness of such actions (e.g., the particular disadvantages that each refugee is facing, the expression of gratitude and sadness of the refugees, etc.). After listening to Jane's testimony, Mary could experience compassion to a great extent and imagine vividly what Jane is telling her, which in turn can lead her to acquire more moral understanding. After all, as Wodak has argued, Mary can find reasons to defer to Jane's moral testimony.[44] Nonetheless, MES provides an explanation as to why Mary will not acquire the *highest* level of moral understanding-why helping war refugees is right only by deferring to Jane. The key idea is that, given the *nature* of emotions, certain first-hand emotional

[42] This act-token—helping war refugees—would be an act-token of a type of generous action. Generous actions are morally appraised as 'right.'

[43] The first-hand emotional experience that I have in mind is also dynamic (Ben-Ze'ev 2010: 57). In my view this entails that, for example, a successful first-hand episode of compassion helping war refugees can lead to understand the rightness of this action to the greatest extent, so Jane has to experience compassion *during* the act of helping war refugees, but she could also experience compassion afterwards. When Jane remembers what it was like to help war refugees, she can keep on acquiring understanding of the right-making features of this action whether she experiences compassion again or not. The point is that Jane could not have acquired the highest level of moral understanding-why helping war refugees is right, if she had not experienced compassion first-hand during the act of helping war refugees. I will start explaining why this is the case later in this section.

[44] Wodak's argument concerns cases of pure moral deference, whereas this example is describing a case of impure moral deference. Nevertheless, my pessimism about moral deference grants that we can find reasons to defer to the testimony of others, in both pure and impure cases of moral deference. My pessimism specifically targets the impossibility of the acquisition of the highest level of moral understanding.

episodes can lead to the acquisition of moral understanding in a distinctive way. I will develop this claim in detail in chapter V. In chapter III, I will properly introduce and discuss the view of emotions that I will be endorsing in this thesis.

Setting emotions momentarily aside, it is easy to see how first-personal experiences are important to moral understanding. Sliwa says:

> 'We often say that only by seeing something first-hand is how "we really got it" or "it finally clicked." What is it that firsthand experience gets us? Actions are right and wrong in virtue of their features. To understand why an action is wrong generally requires you to know what (some of) its wrong-making features are. First-personal experience gives you a richer conception of what those are. Contrast being told that a patch of color is red with looking at it. In both cases you may come to know that it's red. But in the latter case you learn a lot more: you learn that it's red by seeing its precise shade. Similarly, compare being told that prisons are dehumanizing with visiting a prison yourself. In the latter case, you learn a lot more: you come to know that it's dehumanizing by seeing the myriad ways—big and small—in which prisoners are dehumanized. Of course, you can learn more detail by seeking out more detailed testimony. But, setting aside that we may lack words to express some of what we see, even the most detailed testimonial account cannot rival the richness of the content of our own perception.' (2017: 548)

All things being equal, it is clear that by having first-personal access to certain situations or objects (i.e., perceiving them first-hand) provides you with more information than the most detailed testimony, and therefore it is more likely that you understand the object of your perception better. Sliwa accurately describes the way in which first-personal experience can provide us with a 'richer conception' of the rightness and wrongness of actions. The first-personal experience of witnessing a crime of robbery, for example, may be crucial to gain epistemic access to certain wrong-making features of this crime, such as what being the victim of robbery is like. However, undergoing the first-personal experience of *being* the victim of robbery, is possibly the only way in which you can *fully* appreciate all of the wrong-making features of such crime. As Sliwa points out, first-personal experiences cannot only allow us to acquire more moral understanding, but they can also *broaden* our capacity of moral understanding (Ibid: 549). This fact can be explained by the abilities that one can gain by having a particular first-personal experience according to Lewis:

> '…you gain abilities to remember and to imagine. After you taste Vegemite, and you learn what it's like, you can afterward remember the experience you had. By remembering how it once was, you can afterward imagine such an experience. Indeed, even if you eventually forget the occasion itself, you will very likely retain your ability to imagine such an experience. Further, you gain an ability to recognize the same experience if it comes again.' (Lewis 1988: 17)

Hence, being the victim of a robbery can give you *abilities* to remember and imagine what it is like to be robbed, as well as the ability to recognise the same experience when you or someone else are in danger of being robbed. This personal experience can also increase the range of circumstances for which you are in a position to know what the right thing to do is (e.g., how to help a victim of robbery). Sliwa also emphasises that first-personal experiences can lead us to *empathise* with others and put ourselves in the shoes of others in similar circumstances (Sliwa 2017: 549). She also explains that according to her Moral Knowledge Account, '…imagination, moral perception, emotional responses, and reflection are all ways of achieving moral understanding and all have their own distinctive phenomenologies' (Ibid: 550). However, she does not provide an account of the importance of the role that first-personal emotional episodes can have both in broadening our capacity of moral understanding, and in leading us to achieve more instances of moral understanding. In contrast, my MES will attempt to show the way in which emotions can *enhance* first-personal experiences, and therefore contribute to the possible acquisition of a higher level of moral understanding, as well as to broaden our capacity of moral understanding. Although the fact that first-hand emotional experiences can enhance our capacity of moral understanding is part of my MES, I will be mainly concerned with describing the ways in which these experiences can lead us to achieve higher levels of moral understanding.

MES's new epistemic pessimism applies to cases of both pure and impure testimony. According to MES, even if moral testimony can convey true judgements, the degree of moral understanding transmitted through moral deference would still not be the highest, regardless of the truth or accuracy of the moral testimony in question. Also, unlike moral pessimism, I will not deny that it is possible to acquire moral knowledge and moral understanding that can lead to morally worthy actions via moral deference.

The main claim of epistemic pessimism is grounded by the fact that by directing our *attention* to what is relevant in a moral situation, emotions can provide *the most detailed picture of the moral features at stake*. Therefore, a direct experience with moral facts devoid of emotion would not involve highest moral understanding.

In the next subsection, I will briefly sketch the role that the attention provided by first-hand emotional experiences can play in the acquisition of moral understanding.

iii.i Attention Provided by Emotion. A First Glimpse.

Although in chapter IV I will explore in detail the way in which emotions seem to direct our attention, here I will introduce the relevance that this feature of emotion can play in acquiring the highest level of moral understanding.

Some authors have argued that emotions involve patterns of salience (de Sousa 1987; Gibbard 1990; Elgin 2008; Ben-Ze'ev 2010; Tappolet 2016). In other words, they have argued that different emotions focus our attention on different things. Tappolet says:

> '…it is important to keep in mind that what appears to be true of one kind of emotion is not necessarily true of others. Consider the relation between emotion and attention, for instance. It is plausible that when one experiences fear, attention is focused on what one is afraid of. But emotions such as joy or boredom appear to have a very different influence on attention. When you are bored at a concert, your attention drifts away from the music as you start thinking about some philosophical puzzle, say. As experimental work suggests, joy and more generally positive affective states come with a widening of attentional focus.' (2016: 5)

Due to these differences in focus of attention and other facts about emotions, I will later discuss how some emotions can negatively affect the first-hand emotional experience that I have in mind.

For brevity's sake, in this chapter I will only describe the case of a successful first-hand emotional experience, tied to an example of moral deference. A famous positive account of the way in which emotions can direct attention is Michael Brady's (2013). Brady has argued that emotions can promote the understanding of our evaluative judgements, by *capturing and consuming* our attention (Ibid:158). According to Brady, acquiring understanding of a given situation through emotion involves a search for reasons that enables us to assess our initial judgement of it, more accurately. For example, through the experience of fear, there is usually an appraisal of some threat. In cases where fear does not trigger an immediate—or as it is often said, instinctive— reaction, fear seems to fix our attention on the threat; before deciding on an adequate response to it, according to Brady, fear would involve a search for reasons why some object or event is or is not dangerous. These reasons would be considerations which allow us to understand the seeming dangerousness that was initially appraised by fear. Therefore, upon reflection caused by fear, I can arrive to the conclusion that the dangerousness of speaking in public, is not appropriately understood by my initial response, say, crying.

Although Brady's account illustrates one way in which emotions lead to *evaluative* understanding by directing our attention, in the rest of the chapters of this thesis I will provide a more detailed view of the role that emotional episodes can play in the acquisition of *moral* understanding by directing our attention (especially in chapter IV which is solely dedicated to this topic). For now, I will provide an example of the impossibility of acquiring the highest level of moral understanding via moral deference in the next subsection.

iii.ii First-hand Emotional Experiences and Moral Deference

Here is Mary's first-hand emotional experience outside the laboratory, witnessing a case of bullying (an act-token of a cruel type of action, morally appraised as *wrong*):

(A) Mary sees John bullying Bob, and all things being equal, she experiences *compassion* for Bob. Compassion fixes her attention on the way John is bullying Bob, the fact that Bob is crying, and that John is laughing hysterically at John's misery. Bob's tears and John's mocking, provide her with good reasons for her to think that John is doing a morally bad thing. The memory of Bob's and John's reactions leads her to imagine what both of them must have felt like. She judges that John did a morally bad thing.

Compare it with a case of pure, moral deference:

(B) Bob tells Mary: "John acted badly towards me." Mary defers to Bob, and she believes that John did a morally bad thing.

Mary feels *compassion* because Bob is her close friend, though she isn't provided with further reasons why John did a morally bad thing. Therefore, she only knows and understands to a low extent that John did a morally bad thing. Nonetheless, since she has some reason (e.g., Bob is generally trustworthy) to trust Bob's testimony. Her compassion can also lead her to imagine what it was like for Bob to be wronged. She defers to Bob, and judges that John did a morally bad thing.

Compare with a case of impure moral deference:

(C) Bob tells Mary: "John acted badly towards me. He was mocking the way I talk, the way I dress, and he ridiculed me in front of everybody. I told him a very personal secret a while ago, and he started to discuss it with everyone out loud."

Mary feels *compassion* by listening to Bob's testimony. She understands to a high extent why John did a morally bad thing. She defers to Bob and judges that John did a morally bad thing.

If we compare (B) and (C) with (A), it is possible to see that although there can be an emotional engagement via cases of pure and impure moral deference, the first-hand emotional engagement in (A) is more epistemically rich for Mary, in the sense that it allowed her to *focus her attention* on the wrong-making features of John's action towards Bob, and therefore provided her with the information necessary to better understand why what John did to Bob was wrong. Now compare (A) with (D):

(D) John bullies Mary, and she experiences *indignation.* Shame fixes her attention on the way that John is mocking her, and on the way in which his mocking makes her feel uncomfortable. Shame directs Mary's attention to John's cruel laughter and look of contempt. These features

> of John's action provide her with good reasons to think that John has done a morally bad thing. The memory of how John's insults made her feel, leads her to further reflect on the reasons why what John did was wrong. The memory of the way in which John's insults made her feel, provides her with the ability to recognise future cases of bullying and the desire to stop them. She judges that John did a morally bad thing, and understands to the highest extent why bullying is wrong.

Both (A) and (D) show different kinds of first-hand emotional experiences. In (A), Mary is a mere spectator undergoing an emotion, whereas in (D) she is the person being wronged undergoing an emotion. In both cases, Mary's emotional episode of compassion and shame directed her attention to the morally relevant features of what John's act of bullying. However, although the degree of moral understanding that she acquired in (A) was very high, the degree of moral understanding that she acquired in (D) was the highest in the sense that for Mary to acquire the *very highest* level of moral understanding of the wrongness of bullying, she would have to emotionally experience all of the possible instances of bullying. In other words, she would have to be the victim of *all* types of bullying, and plausibly the perpetrator of all these types of bullying with the corresponding emotional episodes (e.g., shame, anger). Which seems difficult and undesirable.

In these cases, I have assumed that Mary's episodes of compassion and shame did not misfire, but they easily could have directed Mary's attention to the morally irrelevant features of John's action. For example, compassion for Bob could have directed her attention to someone else, and not notice the relevant wrong-making features of John's bullying. Alternatively, shame could have focused her attention solely on her defects, and therefore come to judge John's bullying as 'right.' In chapter III I will discuss some of the ways in which emotions can misfire (e.g., Goldie 2008), although in chapter V, I will also describe a kind of 'virtuous emotional acquaintance' which would warrant reliable first-hand moral emotional experiences.

Finally, I will say that if we compare cases (A), (B), (C) and (D) with emotionless Mary inside the laboratory, it is hard to see how her attention could have been directed to the morally relevant features of bullying. Even if inside the laboratory she already understood the reasons why bullying was wrong, it is easy to see that the phenomenology of her emotional experience in the cases just described increased her understanding. I will argue for this claim in more detail later in chapter V.

Moral Epistemic Pessimism about moral deference presents a challenge to Wodak's optimistic argument for having reasons for approving, acting, and believing on the basis of pure and direct testimony. Even if these reasons hold, pure (and impure) moral deference would still be problematic in the sense that it will never provide the highest degree of moral understanding.

In the next chapter, I will describe the components of emotions, and discuss some of the ways in which emotions can lead us to acquire knowledge and understanding, as well as some of the ways in which they can lead us astray.

Conclusion

In this chapter, I have argued that moral deference is problematic in the sense that it does not provide moral understanding of the highest level. First, I discussed two views that describe moral deference (pure and impure) as problematic in different ways (i.e., moral pessimism, Fletcher's psychological pessimism). According to moral pessimism in general, deferring to the moral testimony (pure and impure) of others is problematic given that it leads to insufficient moral understanding of the reasons required to perform morally worthy actions. According to Fletcher's psychological pessimism, pure moral deference is problematic since it is difficult for the receiver of the testimony to form the moral sentiments associated with or that constitute the moral judgement of the giver.

Second, I discussed Wodak's view, according to which there always seems to be at least some reason to defer to the pure moral testimony of someone else, and which allows for sufficient moral understanding and morally praiseworthy action, as well as for the formation of moral sentiments. I agreed with Wodak's optimistic claim about the unproblematic nature of pure moral deference, in the sense described by moral pessimism and by Fletcher. Third, I described Callahan's view. According to Callahan's form of moral pessimism, moral deference is problematic given that it leads to insufficient moral understanding (reconceived as an affective and motivational engagement with moral reasons), and that it discourages the receiver to keep on acquiring moral understanding. I then argued that although Callahan's moral pessimism highlights the importance of emotions in the acquisition of moral understanding (as opposed to other kinds of moral pessimism), it still does not provide an account of the most significant epistemic problem with moral deference.

Finally, I described my Moral Epistemic Pessimism about moral deference. In contrast with moral pessimism and Fletcher's psychological pessimism, I argued that moral deference is problematic given that it cannot provide the highest level of moral understanding, due to the receiver's lack of a first-hand emotional experience with the morally relevant features of a certain action.

In the next chapter, I will describe the account of emotion that I will defend in this thesis, and describe some [positive] putative epistemic roles of emotions.

Chapter III

Can Emotions be Epistemically Useful?

My aim in this chapter is to argue that, unless they misfire, emotions can play different positive epistemic roles. By 'epistemic role' I am referring to a part of the process involved in the production or modification of knowledge and/or understanding. Concretely, I will defend the claim that emotions can be sources of salience (e.g., de Sousa 1987; Hookway 2008; Elgin 1996; 2008; Ben-Ze'ev 2010). In other words, emotions can sometimes direct our attention to certain aspects of a given situation in a way that produces or modifies our knowledge and understanding of it. For example, when she encounters them, Mary's fear of spiders makes certain features of spiders salient, as well as the ways in which she can get rid of the spiders. Mary's fear-directed attention can provide her with—sometimes new–information about spiders: where they hide, their different sizes and colours, etc. By noticing their sizes and colours, Mary can investigate which spiders are venomous and which are not. Alternatively, knowing where spiders hide allows Mary to avoid these places (e.g., the attic).

Each emotion directs our attention to different types of features of a given situation (Gibbard 1990), thereby playing the positive epistemic role of motivating the process of acquiring knowledge and understanding. In the next chapter, I will argue that direction of attention will be the most important epistemic feature of emotions for my Moral Epistemic Sentimentalism (MES). Recall that MES is an account of moral understanding that emphasises the importance of the role that emotions can play in the acquisition of moral understanding, and its key claims of MES are (a) instances of moral understanding-why are reducible to instances of moral knowledge (b) moral understanding comes in degrees (c) emotions are sources of salience and can direct our attention to the morally relevant features of morally appraised actions, and (d) emotions can provide moral understanding in a distinctive way given their components. Besides presenting some of the putative positive epistemic roles that emotions can play, this chapter will introduce the ways in which emotions can be sources of salience, and chapter IV will describe in detail the way in which emotions can direct, fix, and consume our attention.

In the first section of this chapter, I will elucidate the theory of emotions that I will be assuming in this thesis (Ben-Ze'ev 2010). My view falls under the category 'hybrid evaluative-feeling' (de Sousa and Scarantino 2021). In contrast with other theories of emotions which belong to different categories (e.g., feeling, motivational), the view that I will develop describes emotions as complex mental states. More precisely, I will argue that emotional episodes are intentional mental states which involve a dynamic process, as well as four components (cognition, evaluation, motivation, feeling). My aim is to develop and defend a view that can explain at least some of the ways in which emotions can lead to the acquisition of different kinds of knowledge and understanding, to later argue that certain emotional episodes can lead to the acquisition of moral understanding.

In the second section, I first describe three positive epistemic roles that emotions may play (i.e., sources of salience, evaluation, and motivational force). I will argue that just as emotions can highlight features of the environment that can increase our understanding of these features (e.g., Mary's fear of spiders can increase Mary's understanding of spiders), emotions can also highlight the features of a situation that are morally relevant (e.g., Mary's guilt for having insulted Bob can increase her understanding of the wrongness of her action). Second, I will describe Catherine Elgin's account of emotions as epistemically helpful (Elgin 2008). Elgin argues that emotions can be sources of salience, and can provide information about the environment (1996, 2008: 41). Although Elgin's account of the role epistemic roles that emotions can play is not a view about acquisition of moral understanding nor of moral knowledge, her claims about emotions being sources of salience and sources of information about the environment can be deployed to explain claims (c) and (d) of MES: emotions can direct our attention to the morally relevant features of morally appraised actions, and emotions can provide moral understanding in a distinctive way given their components. Third, I will discuss the Standard View that characterizes emotions as variable and volatile, and hence epistemically unhelpful (e.g., Elster 1999; Frijda 2010). In response, I will follow Elgin in arguing that although emotions can lead to epistemic error and to irrational or immoral action, emotions can be regulated (e.g., through reflection or refinement through the arts) to ameliorate these negative effects (Elgin 2008: 47-48). Fourth, I will mention some examples of emotions that are typically regarded as epistemic (e.g., curiosity), in order to distinguish them from typically moral emotions (e.g., guilt).

In the third section, I will discuss a distinct way in which emotions might be epistemically useful, namely that suggested by the Perceptual Theory of emotions. The Perceptual Theory holds that emotions are, in essence, *perceptual experiences of evaluative properties* (Tappolet 2016: 15). Theorists of emotion that wish to describe emotions as performing certain epistemic roles (e.g., justification), argue for a view that fits the perceptual model (e.g., Mitchell 2017, Döring 2003, 2008). In contrast, I will argue that emotions can act as perceptions, in the sense that they can allow us *to notice* certain environmental features, although I will not hold that they are in essence perceptions (nor perceptual experiences) of evaluative properties, nor that they justify evaluative judgements. I will also discuss the way in which the Perceptual Theory deals with the recalcitrance objection, which explains that we can experience emotions that we are *not supposed* to experience given certain perceptions or judgements (Grzankowski 2020). For example, Mary's fear can be recalcitrant when she judges that spiders are not dangerous and yet she fears them. Recalcitrance is a way in which emotions can be epistemically unhelpful, so it is a phenomenon that my view will also take into account. Ultimately, emotions can be regulated to avoid recalcitrance, and so it would seem that recalcitrance is not epistemically problematic.

Finally, I discuss Peter Goldie's view of misleading emotions (2008). Goldie argues that although most emotions can be refined and regulated, there are some emotions that can be *systematically misleading*,

due to an environmental mismatch (2008: 155). For example, *fear and mistrust of strangers* used to play a significant adaptive role a long time ago, but today it mainly creates further social problems (e.g., discriminatory behaviour can lead to violence). I argue that Goldie is right in pointing out that emotions can mislead us in a systematic way, and that this fact makes it harder to regulate them. I argue that Goldie's view of the way in which emotions can epistemically misleading is healthy, in the sense that a moderate scepticism towards the roles that emotions may play in our thinking mechanisms will also allow us to define these roles better, and optimise them and use them to our advantage.

i. Theories of Emotion

The account of moral epistemology that I develop in this thesis, requires a specific understanding of the nature of emotions, and a description of certain emotional episodes. Several, but perhaps not every view of emotion that there is, could be accommodated by my view on moral and direct emotional experiences, and on my theory of change in moral perspectives. For now, it is important to specify that whenever I am discussing 'emotions' I am referring to *episodes* (i.e., occurrent instances of emotion), rather than to emotional dispositions (i.e., enduring emotional states).

The theory of emotion that I will endorse belongs to the category of *evaluative*, as opposed to the *feeling*, and *motivational* categories. Roughly put, the Feeling Theory of emotions asserts that emotions are merely a type of bodily feeling (not simply 'gut feelings', but psychological states), and that without the feeling component '…our state of mind can hardly be called emotional at all' (James 1890:471, see Price 2015:17). On the other hand, the Motivational category is divided in two versions: phenomenological, and non-phenomenological. The former assumes that emotions are merely feelings of 'action readiness' (Deonna and Teroni 2015), and the latter views emotions as psychological 'causes of states of action readiness which may or may not be felt' (Scarantino 2015). What is distinctive about all these theories is that they only focus on one feature of emotions. Feeling theories consider that the physiological aspect of emotions and the way they feel is their essential characteristic, and the Motivational theories regard the behaviour they prompt as their key feature. For this reason, these theories seem insufficient for the purposes of describing the seeming complexity of emotion and its various components (mental, evaluative, phenomenological, behavioural, for instance). I intend to endorse a view which describes emotions as having various components, given that such an understanding of emotions would best account for both our general emotional experiences, and for at least some of the ways in which we acquire moral understanding. I will develop this idea in detail in chapters V and VI.

There are some distinctions to point out when it comes to describing the *evaluative category* to which some theories of emotion belong to. There are different kinds of evaluative theories of emotion, and at

least one type of them will be adequate for the view of emotional episodes that I have in mind. Each example of the views that I will be briefly discussing below, faces different problems and objections, but for now my aim is only to distinguish some evaluative theories of emotion, in order to later locate my view among them in this section.

The first distinction among evaluative theories of emotion to be aware of, is that some of them are *constitutive*, and some of them *causal* (de Sousa, Scarantino 2021: V). Constitutive views argue that emotions *themselves* are in part particular kinds of cognitions or evaluations, whereas causal views argue that emotions are *caused* by particular kinds of cognitions or evaluations. An example of a *constitutive* view would be that of emotions as judgements. According to this view, at least part of what it is to be in an emotional state is to make a judgement (Grzankowski 2020:501). For example, Mary's *anger* at Bob for forgetting her birthday would be partly constituted by the judgement 'Bob has wronged me' (see Solomon 1973; Nussbaum 2001). Hence, for example, according to judgementalist views of emotion, to fear spiders is (at least in part) to judge that they are dangerous.

In contrast, *causal* views would be identical with 'appraisal theories of emotion' (Arnold 1960; Lazarus 1991; Scherer 2001). Appraisal theories of emotion hold that emotions are produced by some sort of cognitive evaluation such as a judgement, a thought, a perception or an act of imagination (de Sousa and Scarantino 2021: VI).[45] For example, Mary's *anger* at Bob could have been caused by *thoughts* that 'Bob has wronged me.' The aim of appraisal theories of emotion is to describe the structure of the processes by which emotions are elicited (Ibid: VI). According to Arnold, appraisals can be made throughout three primary dimensions: eliciting circumstances can be evaluated as good or bad, present or absent, and easy to attain or avoid (Ibid: VI). For example, a thought that causes fear of getting the virus of Covid-19 can be described as the appraisal of Covid-19 as bad, absent but possible in the future, and hard to avoid.

Arnold has defined the notion of *appraisal* as 'the process through which the significance of a situation for an individual is determined' (1960:171). For example, Bob has forgotten Mary's birthday, and this is significant for Mary in a particular way. The fact that Bob has forgotten her birthday has elicited *anger* in Mary. The way in which Mary *assesses* Bob's forgetfulness can occur in different ways. Mary's anger could be produced, for example, by judging or perceiving Bob's forgetfulness as something negative. Mary's anger could also be elicited by her imagining that Bob went to someone else's birthday party, which she evaluates as something bad. However, when Bob forgot Jane's birthday, Jane felt *relief* instead of anger. In other words, the same fact can generate different emotions in different people. Mary evaluated negatively Bob's action, and Jane evaluated positively, given that the

[45] For Arnold, an emotion is a 'felt tendency toward anything intuitively appraised as good (beneficial), or away from anything intuitively appraised as bad (harmful).' (1960: 171)

individual appraisal (i.e., including personal significance) that they attributed to Bob's action was different.

Roughly, the main two differences between theories of appraisal and of emotions as judgements are (i) that appraisals are *processes* whereby we make a *series* of judgments (i.e., whether the object of our evaluation is good or bad, whether it is present or absent, or easy to attain or avoid) whenever we are in an emotional state, and (ii) in the appraisal theory, emotions are caused by cognitive processes like thoughts or judgements, and in the judgementalist theory emotions are constituted at least in part by the judgement. According to the appraisal theory, emotions are distinct from the cognitive process that causes them, and the evaluation that results from the emotional experience would be the appraisal.

There is another category within the evaluative tradition of emotions labelled as 'hybrid evaluative-feeling' (see de Sousa and Scarantino 2021).[46] This would be the category my view would fit into. There are three main types of views which belong to this category. Two of them are of the constitutive kind, and the third one can be endorsed by both constitutive and appraisal theories. My view would qualify as an appraisal theory of emotion.

Firstly, there is the constitutive theory of emotions as *evaluative perceptions*. These perceptual theories of emotion come in strong and weak forms, and usually involve a hybrid model of perceptions and feelings (Brady 2013). An example of a *strong* perceptual theory is Jesse Prinz's (2004). Prinz's view holds that emotions are perceptions of bodily changes (such as a racing heart, blushing, etc.), which have distinctive functions and specific valence markers (i.e., positive or negative signals), and which motivate action (2004:69). Prinz relies on Antonio Damasio's neuroscientific work to argue that emotions belong to a particular system within our somatosensory system (see Damasio 1994). According to Prinz, emotions allow us to perceive bodily changes, as well as that which the emotion represents (i.e., an external environmental feature). Hence, emotions would work as perceptions of both bodily changes and external features. For example, Mary's anger would involve the perception of the redness coloring her face (through the burning feeling in her cheeks, say), as well as the perception of Bob's offence (which would be eliciting her anger). According to Prinz's theory, Mary's anger would also involve a negative (say) valence signal, and this valence in turn would *motivate* her to avoid Bob. Therefore, in this view emotions are literally perceptions of bodily changes, and these bodily changes also indirectly represent the natural content (as opposed to evaluative) that they are reacting to.

On the other hand, an example of a *weak* perceptual theory of emotion would be Christine Tappolet's (2016) (see Brady 2013; Salmela 2011). Tappolet argues that emotions are perceptual experiences of evaluative properties (2016:15). In other words, emotions represent their object as having specific

[46] The general critique made against the evaluative theory of emotions as judgements, is that it fails to provide an account of the complexity involved in emotional evaluations (e.g., their intentionality, motivational force, phenomenology, etc.). As a result, evaluative and feeling traditions have come together in some respects in order to explain in more detail some features of emotional evaluations.

evaluative properties, such as offensiveness, dangerousness, contemptibility, admirableness, the disgusting, etc. (Ibid: 15). To support this view, she lists analogies between emotional and sensory experiences, some of which are: both are *conscious states* that are characterized by phenomenal properties. For example, there is a way it is like to see something as blue, just as there is a way it is like to experience anger or disgust. Both emotional and sensory experiences are *automatic*, since they respond to the world differently from voluntary action. For example, you can neither decide to experience anger when you do not happen to experience this emotion, nor choose to see white snow as orange, say. Both emotions and sensory perceptions are commonly taken to have c*orrectness conditions* i.e., they can be assessed in terms of their appropriateness (see D'Arms and Jacobson 2000). For example, it is possible to criticize Mary's anger towards Bob's offence, if it does not seem that Bob has wronged her; anger appears to have correctness conditions in the same way as the visual experience of poppies as 'blue' would have correctness conditions) (Tappolet 2016: 19-25). I will describe Tappolet's view in more detail in section iii of this chapter.

The second kind of hybrid evaluative-feeling (constitutive) theory, would be the theory of emotions as *evaluative feelings* (Goldie 2000; Döring 2007; Helm 2009). Peter Goldie, for example, has argued that an emotional evaluation is a 'feeling towards' the object of the emotion (2000:72-83). As opposed to Prinz's theory, by a 'feeling towards' Goldie does not mean a *bodily feeling.* Rather, Goldie is referring to a particular way of experiencing or thinking of an object or situation. For example, when Mary says that she is experiencing anger towards Bob's offensive action, to describe her anger as a feeling towards this action is to imply that it has a distinctive phenomenology: there is something it is like to experience Bob's action angrily. Besides having a distinctive phenomenology, 'feelings towards' are intentional states: they represent objects and situations in certain ways (e.g., as offensive, dangerous, disgusting, contemptible, admirable, etc.).

Even though Goldie takes feelings to be more like perceptions (than judgements or beliefs), he denies that they are literally perceptions; 'feelings towards' are a specific class of intentional state (2009:237-8, see also Price 2015:29). Bennet Helm (2009) argues for a similar view, describing emotions as 'intentional feelings of import' which are either pleasant or unpleasant. Something has an import if one cares about it (i.e., if it is worthy of attention and action) (2009:252).

Finally, in the next subsection I will describe the *patterns of salience* view (de Sousa, Scarantino 2021: 7.3), which explains that every emotion tends to make certain things appear as salient. This view could be endorsed by either constitutive or appraisal theories of emotion. The patterns of salience view will help to elucidate the mechanism whereby emotions can direct our attention to specific environmental features, which is one of the main claims of my Moral Epistemic Sentimentalism.

i.i Patterns of Salience View

The view that emotions can include *patterns of salience* may be endorsed by both constitutive and causal theories (see Schroeder 2007:156[47]; Gibbard 1990:136). Roughly, emotions sometimes are described as performing the role of controlling salience among many perceived objects (de Sousa 1987; Elgin 2008; Ben-Ze'ev 2010). Emotions are said to fix patterns of attention, highlighting certain aspects of a situation and obscuring others (Elgin 2008:43). In other words, emotions can direct attention by revealing aspects thereby considered to be worthy of notice, and therefore relevant to the evaluative judgement in question. For example, anger disposes Mary to look at a situation in ways that would reveal evidence of Bob's offence, and of Mary's judgement that he has wronged her (e.g., his erratic behavior or the fact that he is ignoring her), and opportunities to act in consequence (e.g., by accusing Bob, by avoiding him, or by yelling at him). From now on, whenever I describe emotions as including patterns of salience, I will be mainly referring to the fact that emotion. s can *direct our attention* to specific environmental features.

Hence, anger directs the attention to the most salient features of its object. Mary's angry response highlights Bob's behaviour as salient or worthy of notice. What determines, however, what is considered to be worthy of notice in itself? There are, at least two answers to this question. Firstly, it appears that *something should be salient* in certain circumstances, if one would want or should want to notice it (i.e., have epistemic access to) in those circumstances (Elgin 2008: 25). 'Salient aspects' would be those things we wish we would have known, were we later to find out about them. Roughly, we wish we would have known these things, if noticing them would have involved understanding our situation better. For example, after having an argument with Bob, Mary might wish she had noticed that Bob was under a lot stress due to a high amount of work, and that is the reason why he forgot her birthday. Also, it could be argued that we wish we would have known these things if noticing them would have involved having more justified true beliefs. For example, Mary might wish she had not believed her imagining that Bob forgot her birthday because he went to someone else's birthday party. In other words, the salient aspects of a situation seem to be equivalent to what would be regarded as *useful information* for the formation of an evaluative judgement from the agent's point of view.

Secondly, emotions usually arise when positive or negative significant changes are perceived in our personal situation, or in the situation of those who are related to us (Ben-Ze'ev 2010:42). For example, a promotion in her job elicits Mary's *joy,* and Bob falling from a cliff elicits Mary's *anguish*. Emotions give significance to the events that elicit them, given that events (or the changes that cause them) are evaluated as relevant to our *personal concerns* (Nussbaum 2001). In other words, if Mary did not care

[47] In chapter IV, I will discuss how Schroeder's (2007) account of directed-attention desires is analogous to the way in which typically moral emotions involve patterns of salience.

about her professional career or a about Bob, she would not have experienced joy for her career nor fear for Bob. Ben-Ze'ev's widely known metaphor illustrates that emotions arise when we perceive a significant change in our situation: 'Like burglar alarms going off when an intruder appears, emotions signal that something needs our attention. When no attention is needed, the signalling system can be switched off. We respond to the unusual by paying attention to it' (2010: 42).

However, Ben-Ze'ev also acknowledges that many of our emotions are concerned with what we remember or imagine (i.e., with our mental life). Therefore, there is no need for an actual physical event to occur in order for emotions to arise. Nonetheless, there needs to be a change in our personal psychological environment for an emotion to be elicited (e.g., to remember the past, to imagine the future, to think of what might have been, etc.). As Ben-Ze'ev says: 'The past and the future are part of our present psychological situation and hence changes that might have occurred or that may still occur in past and future circumstances are highly relevant to our present concerns; hence they have a significant emotional impact on us' (Ibid: 43). Accordingly, whether they are elicited by physical or psychological changes, emotions will make salient aspects or features related to whatever it is that we *care* about.

It is important to keep in mind that types of evaluation or appraisal vary, depending on the emotion that is felt. In other words, emotions involve *different* patterns of salience (Gibbard 1990:136). For example, Mary is someone who is easily disposed to feel *anger*. In other words, she is disposed to notice things that actualize her propensity to feel angry. Bob forgets her birthday, and this elicits her anger. Mary's anger then makes salient or striking certain features of what Bob did to her (perhaps Bob's careless manner and the lack of presents were crucial elements that made her feel angry). However, if Mary had been disposed to feel *guilt* (which can be a form of anger at oneself) (Ibid: 126), she probably would have noticed features of what *she* has done in the past instead (perhaps the way in which she always yells at Bob, and continuously recapitulates the times in which he shows carelessness).

Therefore, an important condition for emotions to make certain environmental features salient would be to be *disposed* to experience the corresponding emotions. As seen in the previous example, the patterns of salience that each emotion involves rests on the object of their evaluation: anger can be both other-directed and self-directed. *Shame* works in a similar way. For example, experiencing shame appears to be usually about things that are public (i.e., that concern the judgment of others); however, since it is still possible to feel shame about something that one secretly judges to be inadequate, perhaps this is what is distinctive about shame: it makes salient some personal inadequacy (Ibid: 137).

Now that I have described some characteristics of emotions according to different evaluative theories, in the next section I will describe emotional episodes according to my view.

i.ii Emotions According to Moral Epistemic Sentimentalism (MES)

According to MES, emotions have the following six characteristics:

First, the *intentionality*[48] of emotions. Emotions seem to be reactions to, or *about*, something. In other words, they express subject-object relations (e.g., Mary's anger *at* Bob's action, or love *for* Bob).

The second feature of emotions according to MES would be that emotions involve *cognition through patterns of salience*. In other words, emotions can gather information about the circumstances and so increase understanding. For example, Mary's anger can *direct her attention* to the ways in which Bob is acting strange, and this can lead her to understand Bob's behavior to be mildly or strongly offensive. Hence, the fact that emotions include patterns of salience can allow us to detect specific features of a given situation.

Thirdly, emotions entail a *personal evaluation* of what they are reacting to. For example, by experiencing anger at Bob, Mary assesses Bob's actions as negative, and by experiencing love she assesses them as positive. Mary has a personal concern for Bob and that is why she is not indifferent to his behaviour. The example only shows how the personal evaluations entailed by emotions are positive or negative, in a way that relates to one's own concerns. However, once something is salient or relevant to us, it can be evaluated as positive or negative, even if it bears no direct relation to us (e.g., a dictatorship taking place in a foreign country). When evaluating something, one is never indifferent to the object of evaluation but the concern can vary from one case to another.

Fourthly, emotions involve a disposition or readiness to act appropriately, i.e., *motivation*. This readiness or desire to act is related to the evaluation. For example, if Mary's assessment of Bob's action is negative—she is experiencing anger—, then she will be disposed to act in an antagonistic manner by avoiding him or challenging him.

A fifth important feature of emotions is their *feeling* component, which mainly involves a distinctive phenomenology (Ben-Ze'ev 2010:49). For example, Mary's anger at Bob might involve a headache or a particular state of uneasiness.

Finally, it is worth noting that emotional episodes as described by the above characteristics involve a *dynamic* process, where many elements are at play at different stages. In this sense, it is possible to understand this kind of process as one compatible with appraisal theories of emotion. In such a process, emotions do not constitute the evaluations themselves but rather they are caused by a judgement, some perception, or a thought. Emotions can both sometimes constitute (or cause) the evaluation of an object

[48] This characteristic is worth mentioning mainly to distinguish emotions from other affective states, such as moods and character traits.

or a situation, and also be a part of a process where the evaluation is caused by something else (e.g., a memory).

Given the characteristics provided above, my understanding of emotions throughout this thesis would be most compatible with Aaron Ben-Ze'ev's general definition of emotion (2010:57): 'An emotion is a general mode (or style) of the mental system. A general mental mode includes various mental elements and expresses a dynamic functioning arrangement of the mental system (…) This mode involves cognition, evaluation, motivation, and feeling.'

In sum, the view on emotions according to MES would belong to the hybrid evaluative-feeling category, and it would be of the appraisal rather than of the constitutive kind. Importantly, MES is highly compatible with the patterns of salience view. However, the account of emotion according to MES is more plausible than the alternative views, given that it is multi-faceted and therefore provides a more thorough description of the features of emotions.

After having specified what might be involved in an emotional episode in this first section, in the following section I will describe some of the *epistemic* roles that have been attributed to emotions. I plan to argue that even though emotions might play several epistemic roles, the key epistemic role of emotions that I will defend, is the fact that emotions can direct our attention to the relevant moral features of a given situation. Although perhaps several of the epistemic roles of emotions that some authors describe can be deployed in the moral domain (or in moral epistemology), I will only focus on the capacity that emotions have of directing our attention. This capacity grounds the two main claims of my moral epistemic sentimentalism (MES): (i) certain emotional episodes can lead us to acquire the highest level of moral understanding, and (ii) certain emotional episodes can change our moral perspectives. I will fully develop (i) in chapter V, and (ii) in chapter VI.

ii. Can Emotions be Epistemically Useful?

In this section, I will briefly describe three positive epistemic roles emotions may play (whilst not ruling out that there may be others). My aim in briefly describing these three roles to explore some ways in which emotions can allegedly aid our understanding. In subsection ii.i, I will describe two additional ways in which emotions seem to aid our understanding in general, suggested by Catherine Elgin (2008). In subsection ii.ii, I will discuss two objections to the general view that emotions can be epistemically helpful. In subsequent chapters, I will be endorsing the view that the most useful positive epistemic role that emotions can play when it comes to the acquisition of *moral understanding* is directing our attention to the morally relevant features of a given situation. Although MES will mainly present a positive epistemic account of emotions, it will also acknowledge that emotions can systematically lead us to

error, impair our acquisition of moral understanding, and eventually lead us to act wrongly, as will be discussed in this and the following sections of this chapter.

As seen in the previous section, some theorists of emotion take emotions to involve patterns of salience. Similarly, some epistemologists consider that emotions can function as epistemic *sources of salience*. This role has been emphasised by several authors such as: de Sousa 1987; Lance and Tanesini 2004; Hookway 2008; Ben-Ze'ev 2010; and Elgin 1996, 2008. Roughly, emotions can be sources of salience in the sense that they provide a focus on certain aspects of a situation. In other words, emotions can act as 'spotlights' (Peters 2006:458). For example, fear of spiders can focus our attention on (or spotlight) the objects that might help us to get rid of them were they to appear at our workplace. However, Elgin explains that this epistemic feature of emotions is not as easily described given that this direction of focus involves feelings, attitudes, actions, and circumstances (Elgin 1996:148, 2008: 43). For example, when a hysterical neighbour hears the next door's child crying, they may perceive it as a loud nuisance, whereas the child's parents who experience anguish may hear a specific kind of pain which demands that they draw their attention to ways of bringing relief (Ibid: 153). This example suggests that the same feature—in this case, a cry—can be salient and evaluated in different ways. Hence, an emotion is a source of salience in the sense that it can provide focus on certain aspects of the situation, as well as a way in which that feature is salient (i.e., an evaluation).

Besides considering something to be salient, a constitutive part of considering something to be relevant in a given context is to *evaluate* it (Brun and Kuenzle 2008: 18). Such evaluation through emotions would consist in regarding something as worthy of further consideration. For example, spiders at my workplace are made salient by my experience of *fear* given that I evaluate them as dangerous, and so I need to keep on thinking of ways to deal with them or with my disproportionate fear. Of course, these evaluations can be mistaken due to the fact that we can wrongly find something as salient or relevant. For example, if there are not really that many spiders at my workplace or if spiders are not poisonous, then it could be argued that my fear is disproportionate and therefore my focus and evaluation of spiders as dangerous would be mistaken.

Another epistemic function attributed to emotions is their *motivational force*. The epistemic motivational force ascribed to emotions can be described as emotions motivating cognitive activities. For example, *curiosity* can lead to inquiry, and *frustration* with achieved results can cause one to keep on trying to achieve the desired results. Similarly, emotions can work as mechanisms that appear when our knowledge seems inadequate or not useful (Brun and Kuenzle 2008: 16). For example, *doubt* experienced with *anxiety* can be interpreted as an emotion that motivates critical reflection of the reliability of the outcomes of our research (see Hookway 2008).

Hence, the fact that emotions can be sources of salience, involve evaluations, and have motivational force makes them *prima facie* suitable for performing some positive epistemic functions. Note that the

three epistemic roles of emotions described above are compatible with my account of emotions as involving patterns of salience, evaluation and motivation (as well as intentionality and feeling). However, in the next subsection, I will discuss two ways in which emotions can misfire and be epistemically unhelpful (i.e., volatility and variability). In subsection ii.ii, I will describe Elgin's account of emotional regulation. My aim will be to follow Elgin in arguing that emotions can be epistemically helpful despite their volatility and variability, although I will also distinguish my view from hers. Given that my view will concern particular cases of acquisition of moral understanding, I will defend the claim that the most useful epistemic role that emotions can play according to MES is directing our attention to the morally relevant features of specific situations.

ii.i The Standard View

According to Elgin, there is a popular view that there are at least two features of emotions that seem to render them unfit for adequately carrying out epistemic functions (Elgin 2008: 33). These features are *variability* and *volatility*. The variability of emotions is illustrated by the fact that a single situation can have different emotional responses (Ibid: 39). In other words, different respondents are usually sensitive to different things. To use the example in section ii, the annoyed neighbours might experience exasperation towards the baby's cry, whilst the baby's parents might experience anxiety when they hear that their baby is crying. Also, the emotions experienced over time by the same person can also vary. For example, Bob's anger towards the new elected president might turn to admiration in a short period of time. On the other hand, volatility refers to the fact that emotions can be triggered by pretty much anything (e.g., *fury* can be triggered by a traffic jam or by someone walking very slowly in front of us when we are in a hurry) (Ibid:39).

Hence, some argue that these two features of emotions explain why to be under the sway of emotion is to be *irrational*. Call this the Standard View (SV) (Ibid: 33). According to the SV, emotions impair our understanding, and distort our reason (Ibid: 33-35). Put simply, the SV argues that the variability and volatility of emotions deem them *cognitively untrustworthy*. In other words, we *cannot rely* on them whenever they represent a situation as dangerous, or admirable or contemptible, for example.

There are several examples of claims made by different authors which coincide with the SV. For example, emotions can be considered to be a source of *weakness of the will* (e.g., deciding to do the wrong thing out of jealousy) (Davidson 1970), they can cause actions that do not have *any apparent reason* (e.g., jumping up and down out of joy) (Hursthouse 1991), and they can lead to *impulsivity* (i.e., acting quickly without considering all relevant information) (Elster 1999; Frijda 2010). These and other views seem to be correct in their assessment of regarding emotions as epistemically unhelpful at least

in two ways: (i) they can hinder the formation of true beliefs and evaluations, and (ii) they can motivate irrational and/or immoral action due to the formation of false beliefs.[49] Nonetheless, in the next subsection I will discuss some ways in which some emotional episodes can be regulated and, in some cases, advance our understanding.

Elgin presents d two strategies to argue that emotional representations can be reliably correlated to the event or object that they represent (i.e., the biological and the response-dependence strategies). However, as mentioned above, supporters of the Standard View (SV) claim that emotions are volatile, and therefore epistemically unreliable. Still, it is worth noting that not all emotions appear to be like this. For example, as Hume has pointed out, there are calm passions like *fondness* which persists throughout the time ([1739/40], 276), and even violent passions are not always volatile (e.g., *rancour* can be held for a lifetime). So, if instability or volatility were epistemic impediments, this problem would only be present in the case of some and not all emotions. Moreover, if it is possible to correlate emotions to circumstances in such a way that they provide information about the circumstances, then it is not clear that the volatility feature of some emotions is always epistemically problematic. An example would be the case of *relief*. Relief is usually experienced momentarily, and it can be epistemically helpful in the sense that it can lead to the understanding of having been in worried state (e.g., Mary did not realise that she used to be worried about the result of her exam before she experienced relief when she saw her good mark). More precisely, if it is true that Mary achieved a good mark and she momentarily experiences relief, it is hard to see how in this case this volatile emotion is epistemically problematic.

The SV also holds that emotional responses seem to change greatly from one person to another (e.g., something what makes Mary sad may amuse Bob, and make Jane angry). Given this variability, it seems dubious that emotional representations 'delivered' by these responses can actually provide *reliable* information (i.e., interpersonally stable information) about objects or circumstances. Emotional responses seem to vary according to the different perspectives of different observers, and therefore the content of emotional representations will not be about the nature of the objects nor the circumstances. However, as Elgin notes, the diversity of emotional responses towards a single situation can be explained by the fact that different observers are sensitive to different features of it. For example, Mary sees Jane falls from her bike and feels anxiety, whilst Bob finds the event amusing. Elgin argues that these different reactions would not show that emotions are epistemically unhelpful. The reliability of different emotional responses can be improved if it is regulated along three dimensions: the perspective the subject adopts, how sensitive a subject is, and which emotions dominate in the subject (Ibid: 39). For example, Mary could become aware of the fact that she is overly protective of Jane, and this makes

[49] Richard Joyce says 'They are notorious for lying outside voluntary control; they often ignore our attempts to reason with them; they respond seemingly of their own accord to particular types of stimuli; they affect motivations and thus behaviour.' (2000: 95)

her worry in a disproportionate way whenever Jane finds herself in an awkward situation. Mary could make an effort to realise that when Jane is not in serious danger, she does not have to experience acute anxiety on her behalf.

In the next subsection, I will discuss more ways in which emotional responses can be regulated, and hence in some cases provide reliable information about objects or circumstances.

ii.ii Emotional Regulation

As mentioned above, it is possible to correct and regulate emotions just like in cases of perception (Elgin 2008: 40). For example, Bob notices that the sun is going dark during a solar eclipse, and tells Mary to look at it. However, Mary is not wearing her glasses so she cannot witness the eclipse, and replies to Bob that the sun is not going dark. When Bob tells Mary to put on her glasses, she finally sees the sun going dark, thereby correcting her initial perception. Analogously, although Mary might get overly anxious every time Jane decides to ride her bike, with time she might find soothing techniques (e.g., breathing exercises) that help her to keep her anxiety in check, and realise that riding a bike is not as dangerous for Jane.

According to Elgin there are at least two strategies to regulate our emotions. The first one consists in attending to and reflecting on the following aspects of our emotional responses: the emotions themselves, the situations that trigger them, the orientations they give rise to, and the opinions that they generate (Ibid: 47). This way we can develop either more nuanced or more accurate responses; in other words, this is a way in which we can *refine* our emotional responses.

For example, recall that Mary is easily disposed to feel anger. If Bob happened to forget her birthday again, she might refine her angry response by (i) noticing the unpleasant character of the experience of anger (ii) noticing that her anger is usually triggered by the fact that Bob forgets special occasions (iii) noticing that her anger orients her to pay close attention to Bob's social clumsiness and (iv) noticing that other people tell her she is too harsh on Bob whenever he forgets either her birthday and/or special occasions. Paying attention to this features of her emotional response might lead her to reflect that (i) given the unpleasant experience that comes with it, she would rather not undergo anger (ii) Bob tends to forget other people's birthdays, so it is not worth taking special offence and getting angry at the fact that he forgot her birthday (iii) Bob has many other virtues despite his social clumsiness and her anger blinds her from this fact and (iv) given other people's opinion on her anger towards Bob she realises that her anger might be disproportionate. Ideally, this reflection would lead Mary to appease her anger and not be as upset whenever Bob forgets her birthday or any other special occasion, given that this

reflection reveals her anger to be not fitting. In other words, Mary's anger is not presenting Bob's action with accurate evaluative features (D'Arms and Jacobson 2000: 72).

Hence, refining our emotional responses can involve distinguishing at least four aspects of their occurrence, and reflecting on different aspects of them.

The second (and less discussed) strategy Elgin has suggested to regulate emotions (and therefore improve their epistemic yield) is through *calibration* (2008: 47). Calibration is a particular kind of regulation in that involves further self-knowledge. The self-knowledge that concerns this kind of regulation consists in assessing our degrees of sensitivity to evaluative properties (e.g., shamefulness) (Ibid:47). In other words, our emotional responses can be graded as high or low. According to Elgin, paying attention to the intensity of our emotional responses will also involve discovering which emotions are epistemically helpful and which are not. For example, by realising that the intensity of her angry responses is high, Mary might realise that her anger is not useful for identifying the right reasons why Bob tends to forget special occasions and discovering this fact would itself be an improvement in her understanding of Bob's social clumsiness.

Although some emotions might not seem epistemically helpful, Elgin argues that it is possible to improve them through the arts (Ibid: 47-48). She claims that we often engage with works of art that elicit negative emotions. For example, horror films and dramatic plays that elicit extreme fear and/or anguish, given that they allow us to experience these emotions in muted forms and to explore the perspectives they might yield (e.g., if we are lucky, we will never directly experience the horror of realising that we have murdered our father and married our mother like King Oedipus did) (Ibid: 48). Elgin says:

> Imagining your way into Oedipus' horror, adopting the perspective it provides, and seeing how the world looks from that perspective enriches your life. For you gain the ability to see and feel and discern and respond in ways you previously could not. If emotions afford epistemic access to things, and the arts refine, extend, heighten, and provide opportunities to experience emotions, the arts contribute significantly to cognition. And if we ask why we enjoy tragedies, horror films, and other art forms that elicit negative emotions, the answer is, at least in part, because we enjoy expanding and exercising our abilities, as engagement with such art forms enables us to do. (Ibid: 48)

Alternatively, paying attention to the emotions of artists would be another way in which emotional responses can be calibrated. For example, delight in music makes a director of an orchestra pay attention to the right times when the different instruments should play. In chapter V, I will discuss in more detail how imitating the emotional responses of supposed moral experts would be another way in which emotions can be both refined and calibrated. Elgin's strategies for emotional regulation seem to require effort and presuppose and demand that we become especially interested in the arts. However, studying

and reflecting on our emotional responses, as well as turning to art seem like plausible and useful ways to educate our emotions.

In sum, according to Elgin, emotions can be epistemically useful in the sense that they can be sources of salience, provide information about the environment, and epistemic access to response-dependent properties. I follow Elgin in arguing that emotions can be sources of salience, and can provide information about the environment. I also argue that they can be evaluative and can provide motivational force for inquiry. In principle, Elgin would also agree that emotions can be epistemically useful in these two ways. The main differences between Elgin's view and mine are that (i) I remain neutral about gaining epistemic access to response-dependent properties, and (ii) I am mainly concerned with emotions aiding acquisition of moral understanding, and I will stress the importance of emotions including patterns of salience.

Lastly, as the SV points out, emotions can also be variable and volatile. However, although emotions can misfire in this and other ways, they can also be regulated and corrected. For instance, it is possible to notice our degree of sensitivity, and to identify which emotions dominate our life. We can also regulate our emotions by listening to the advice of others, attending to and reflecting on the situations that trigger the emotions, and to the opinions that they generate. We can also calibrate our emotions by getting to know how sensitive we are, and we can educate them, for example, by turning to the arts. In the next subsection, I will briefly mention some examples of epistemic emotions, in order to later distinguish them from the moral emotions.

ii.iii Which are the Epistemic Emotions?

Various authors have argued that there are emotions which are tied to certain epistemic contexts and therefore have special epistemic character (e.g., Scheffler 1977; Hookway 2008; Tanesini 2008). Scheffler claims that *surprise*, *joy of verification* and *disappointment or joy of falsification* are 'cognitive emotions' which presuppose certain mental states (Scheffler 1977: 12). For example, according to Scheffler, surprise involves the 'supposition that what has happened conflicts with prior expectation' (Ibid: 12). Surprise seems to be epistemically relevant given that maintaining a certain receptivity to surprise might avoid attitudes of radical scepticism, radical credulity, dogmatism, and a form of epistemic apathy, all of which can be an obstacle for inquiry (Ibid: 13).[50]

It would seem that although some emotions can play such epistemic roles, this may not be the case for all emotions. However, authors who have argued for the epistemic functions of some emotions have

[50] Markus Wild has argued that unless emotions are regulated, they do not guide inquiry but rather add to epistemic self-questioning (2008: 125).

not specified that these functions should only be attributed to a selected group of emotions (e.g., de Sousa 2008; Elgin 2008; Thagard 2008). For example, de Sousa classifies *fear*, *greed*, *feeling of doubt*, *feeling of certainty*, *feeling of knowing*, and *feeling of familiarity* as 'epistemic feelings', although he also argues that emotions in general can be epistemically relevant to various degrees (2008: 186). According to de Sousa, any emotion that affects conviction, inference, or cognitive strategies more or less directly will count as an epistemic feeling. For example, if we experience trust by deferring to someone's testimony, we will very likely be convinced that their testimony is a justified and true belief (Ibid: 186). Another paradigmatic epistemic emotion is the *feeling of simplicity* (Hookway 2003, 2008).[51]

The previous examples of epistemic emotions show that an epistemic emotion does not only need to be involved in the acquisition of knowledge (i.e., justified and true beliefs). It seems that at least some emotions can also be a part of the process of cognitive inquiry. Hence, any emotion that is involved in this process could in principle be regarded as epistemic. In chapter VI, I will discuss which are the typically *moral* emotions (e.g., guilt, disgust). Roughly put, the main difference between typically epistemic emotions and typically moral emotions—according to my view—would be that moral emotions can constitute, arise from, or be associated with moral judgements (i.e., judgements that deploy certain thick concepts such as right, wrong, honest, generous, etc.). In contrast, epistemic emotions would be necessarily associated to a specific kind of inquiry (e.g., scientific).

In the next section, I will briefly discuss the Perceptual Theory of emotions. It is widely held that emotions are perceptions of evaluative properties (e.g., Tappolet 2012, 2016; Cowan 2016; Mitchell 2017). My motivation for discussing the Perceptual Theory is to distinguish it from my view. Whilst I ascribe to emotions an epistemic role that aids the acquisition of moral understanding, I will argue that emotions can act as perceptions in the sense that they can allow us *to notice* certain environmental features. However, I will not hold that they are in essence perceptions (nor perceptual experiences) of evaluative properties. I discuss these different claims in the following section.

iii. The Perceptual Theory of Emotions and the Relevance of Non-Conceptual Content

The Perceptual Theory holds that emotions are, in essence, *perceptual experiences of evaluative properties* (Tappolet 2016: 15). I will not discuss views that argue that emotions are literally

[51] It is possible to consider the aforementioned feelings and mental phenomena as epistemically valuable and still wonder whether they really are emotions (Brun and Kuenzle 2008: 25). In psychology, for example, surprise and feelings of familiarity are sometimes regarded as 'nonaffective' feelings (i.e., distinct from emotions) (e.g., Bless et al. 2004, Stepper and Strack 1993).

perceptions of value, given that dealing with the objections that they face is out of the scope of this thesis (e.g., Johnston 2001, Prinz, 2004, Döring 2003, 2008). Rather, I will briefly describe Tappolet's version of the Perceptual Theory, given that her characterisation of emotional perception as *non-conceptual* will be very useful to describe a way in which moral understanding can be acquired through emotional experiences. Tappolet defines emotions as a kind of perception on the basis of *analogies* with sensory experiences (2016: 45). However, I will not argue that emotions are literally perceptions, nor that they are analogous to sensory experiences. Ultimately, as I mentioned before, the epistemic feature of emotions that will be crucial for my view is that in some cases emotions serve the function of directing our attention to whatever we identify as morally important (e.g., certain actions or feelings, or features thereof).

Henceforth, when discussing the Perceptual Theory, I will be solely referring to the claim that emotions are a kind of perceptual experience (i.e., Tappolet's view).[52] The main difference between perceptions and perceptual experiences is that perceptions are factive. For example, perceiving a cat as grey entails that it is grey, whereas perceptual experiences are not factive (e.g., perceiving a grey cat as black). Given this important distinction, the Perceptual Theory (PT from now on) holds that emotions are 'perceptions of evaluative properties unless they misfire' (Ibid: 15). For example, unless it misfires, *disgust* would consist in the perception of something as *disgusting*. On this account, emotions are described as having representational content. In other words, they present their object in a specific way and so, the emotion of disgust with respect to a loaf of bread will be correct just in case the bread is really disgusting (e.g., if the bread is mouldy).

An aspect that distinguishes the PT from other evaluative theories of emotion (e.g., Judgemental theories), is their take on the representational content of emotions (i.e., their intentionality). The PT argues that the content represented by emotions does not need to be conceptual (e.g., articulated via structured propositions) (Ibid: 16). So, for example, when experiencing disgust towards mouldy bread, the representation of the bread as disgusting—provided by my experience of disgust—does not need to come in the form of the propositional mental state 'the bread is disgusting.' Another way in which the content represented by emotions is non-conceptual, on this view, is that we do not need to possess evaluative concepts in order to experience emotions (Ibid: 16). For example, it is not necessary to possess the concept of the disgusting to experience disgust. Hence, although emotional perception would be of states of affairs that involve evaluative properties, it would be 'simple' in the sense that it does not necessitate judgements and concepts. I do not consider that being non-conceptual entails being non-epistemic. In the end, different mental states can be epistemic if they are involved in a process of knowledge or understanding acquisition[53] (whether they facilitate or hinder it).

[52] A similar view is found in D'Arms and Jacobson (2010), Goldie (2004), and Anthony Kenny (1963).
[53] As mentioned before, authors such as Elgin (1996) and (Dretske 1989) have argued that beliefs are dependent on the process whereby they are acquired ('justified beliefs' in Elgin's but not in Dretske's view). However, in my

Hence, there are at least two ways in which emotional perceptions can be non-conceptual. In chapters V and VI, I will argue for a view of acquisition of moral understanding that includes both conceptual and non-conceptual content. Roughly put, conceptual content is found in judgements/propositions and non-conceptual content in sensory experiences (Ibid: 16; Peacocke 1989; Bermúdez 1998). Recall that the main claim of the PT—that emotions are perceptual experiences of evaluative properties— rests on an analogy with sensory experiences, and so this will be the explanation as to why emotional perception can be non-conceptual. For example, consider your visual experience of a pyramid that is steep and stony, and your judgement that the same pyramid is steep and stony. The visual experience and the judgement are both about the same pyramid, but they intuitively represent their object and its properties in different ways. The visual experience is like a *picture* of the pyramid whilst the judgement is like a *description* involving terms that ascribe properties to the pyramid. Contrary to the formulation of your judgement, having a visual experience of the steep and stony pyramid does not necessitate the possession of the concepts 'steep' and 'stony.'[54] Additionally, judgements play a special role in reasoning, whereby they form complex (i.e., structured) inferential networks that are associated through the same concepts (Tappolet 2016: 17). For example, to explain the inference that from the judgement that 'this pyramid is *stony*' and the judgement that 'this sculpture is stony' to the judgement that 'at least two monuments are stony', it seems necessary to assume that the content of the beliefs is structured, and that all three involve *stony* as a constituent. To summarise this last distinction, a mental state will be conceptual if it has content that involves concepts as constituents, and it is not conceptual if this is not the case.

According to the PT, although emotions can and often do involve conceptual content (e.g., it is clearly necessary to possess the concept of 'financial meltdown' to experience *fear* that there will be a financial meltdown) the evaluative appraisal or assessment that comes with the content of emotions is non-conceptual (Ibid: 18). In other words, according to proponents of the PT, it is not necessary to possess the concept of the fearsome to undergo fear and thereby to represent something as fearsome, just like simpler creatures can experience fear and yet do not possess the concept of the fearsome.

Tappolet says:

> 'Important epistemological implications follow from this account of emotions. If emotions are non-conceptual representations of evaluative properties, then it should be expected that emotions are like sensory experiences in that they allow us to be aware of certain features of

account of moral epistemology, moral understanding will not only come in the form of justified or reliably acquired belief. Upstream moral understanding can involve abilities to detect right courses of action, as well as emotional (i.e., non-conceptual) perceptions in the way that Tappolet describes them (e.g., Mary would understand that certain actions are *blameworthy* through the perceptual experience caused by *guilt*).

[54] Claims about the possibility of non-conceptual content are controversial. Roughly put, some authors argue that beliefs about the world that we perceive can only be formed via conceptual content. For example, the conceptual content of the perceptual experience of a shade, can be explained by the demonstrative concept of 'that shade' (See McDowell 1994; Brewer 1999).

> the world. Just as the visual experience of a blue mountain allows us to be aware of the color of the mountain, the experience of fear would allow us to be aware of the fearsomeness of things. More precisely, since emotions can misfire, fear would allow us to be aware of fearsomeness under favorable circumstances, when nothing interferes with it. Given this, it appears plausible to claim that evaluative judgments that are grounded in emotion are *prima facie* justified.' (2016: 18)

The claim that emotions serve as justifications of evaluative judgements even by drawing analogies with perceptual experience is controversial, and I will only discuss it briefly in the next subsection. I will remain neutral about this epistemic feature of emotion, although it has been widely defended (e.g., see Mitchell 2017 and Cowan 2018).

Besides their being non-conceptual, there at least other five points of analogy between sensory experiences and emotions: (i) they both are usually consciously experienced states with a given phenomenology (ii) they are automatic (iii) they are world-guided[55] (iv) they have correctness/fittingness conditions (v) they can be non-inferential (Tappolet 2016: 19-24). These five points serve as further reasons to adopt the PT, but I will not discuss them since I do not plan to endorse this view on emotion as perceptual experiences of evaluative properties. If it is correct that, unless they misfire, emotions can function as a kind of perception of evaluative properties then it is yet another way in which emotions can play a positive epistemic role (i.e., they provide epistemic access to evaluative properties).

The point of analogy with sensory experiences that will be compatible with my view, is the widely agreed fact that emotions are *automatic* (see Descartes 1649; Alston 1967; Gordon 1987; Calhoun 1984; Ekman 1999; Deonna and Teroni 2012; Tappolet 2016). Emotions and sensory experiences are taken to be automatic in the sense that they are not directly subject to the will (Tappolet 2016: 19). In other words, they differ from voluntary action due to the fact that they are triggered automatically, in response to the world (Ibid: 19). For example, you can neither *decide* to feel disgust when you do not happen undergo this emotion, nor *choose* to see stony pyramids as sandy. It can be objected that in fact, it is possible to decide how to feel, or choose to see things that are not there at a given time via voluntary imaginings. Also, there seem to be indirect ways to control our emotions, as it was discussed in section ii.iii (e.g., breathing slowly to avoid panic). However, it is still typically the case that if you encounter

[55] Tappolet argues that just like sensory experiences are usually caused by facts or events in the world (e.g., the poppy and its colour are causally responsible for our experience of the poppy as red), emotions are usually caused by facts or events in the world. For example, the crocodile that emerges from the pond can cause us to experience fear (2016: 20). According to Tappolet, although it is true that imagining something can also cause an emotional response (e.g., fear can result from vividly imagining that a crocodile is swimming in one's swimming pool), emotions are still world-guided in the sense that they are responses to how things are in general in our environment. As discussed in the first section of this chapter, Ben-Ze'ev explains that emotions can be caused by our 'psychological environment', so they are not necessarily world-guided in the sense that they are always caused by physical events in our environment (2010: 43).

mouldy bread in your kitchen and you react with disgust, this disgust is not caused by a decision or an intention to feel disgust. Whatever their kind, emotions in general automatically arise in response to the world (Ibid: 21).

I will return to discuss this important feature of emotions in chapter IV. In the next subsection, I will briefly discuss the problem of recalcitrance and the PT's way of dealing with it.

iii.i The Problem of Recalcitrance

In this subsection I briefly discuss the so-called *recalcitrance* of emotions. This is one sense in which emotions are commonly considered to be irrational (e.g., Brady 2007, 2009; Helm 2009). A recalcitrant emotion is one that persists "despite the agent's making a judgement that is in tension with it. A recalcitrant bout of *fea*r, for example, is one where the agent is afraid of something despite believing that it poses little or no danger" (D'Arms and Jacobson 2003: 129, emphasis mine).

These are some examples of cases of recalcitrant emotions: *disgust* despite believing that the bread is not mouldy, *anger* towards one's partner despite believing they have not wronged you, *guilt* despite believing one has done nothing wrong, etc.

Anyone who is undergoing recalcitrant emotions appears to be 'subject to a certain rational requirement' (Brady 2007: 276). There seems to be a pressure to resolve a seeming conflict between persistent emotions and judgements that contradict them. For example, when we judge that the rollercoaster is safe but fear it nevertheless, or when we judge that our partner has done nothing wrong and yet we feel anger towards them, a normative conflict seems to take place: 'one *should not* both be afraid and judge that there is nothing dangerous, one *should not* be angry with one's partner while judging that the partner did nothing wrong.' (Grzankowski 2020:1) These pairs involving conflicting emotions and judgements are considered to be irrational given that they involve a structural inconsistency (Ibid: 1).[56] This normative conflict can also be explained by the critical assessment through which we judge emotions to sometimes be either fitting or unfitting (D'Arms and Jacobson 2003).

Hence, when experiencing recalcitrant emotions, one seems to be required to either modify the emotions or to change the judgement that partially involves them. Recalcitrance is a phenomenon that every theory of emotions has to deal with. In other words, every theory of emotion is required to provide an account of how to resolve the conflict of recalcitrance: 'When one judges that the dog is harmless and yet fears it, one is doing something one *ought not* do' (Grzankowski 2020: 2).

[56] There are discussions of structural rationality focused on belief and action (e.g., Broome, 1999, 2007), but the present discussion of recalcitrance concerns the structural relationship between judgements and emotions.

The PT deals with recalcitrance by holding that there is no problem with perceiving things to be one way whilst judging them to be otherwise. Tappolet argues that unlike emotions, sensory experiences do not seem to be philosophically assessed in terms of rationality (2016: 31). Even when illusory, there does not seem to be a requirement to resolve an irrational conflict in sensory illusions. For example, it is not clear why someone experiencing the Müller-Lyer illusion should either stop seeing the lines as unequal, or change their belief that the lines are unequal, or it does not appear as incoherent to perceive the pencil as not straight whilst judging that, despite appearances, it is straight (Brady 2007: 276; Grzankowski 2020:2). Hence, by making emotional experiences analogous to sensory experiences, the PT would seem to avoid recalcitrance, or at least the need of resolving any conflict of irrationality.

However, given their recalcitrant character, there does appear to be an important difference between emotions and sensory experiences. Tappolet says that, in contrast to the case of sensory perception, it is possible for us to attempt to get rid of inappropriate emotions albeit indirectly, as discussed in section ii.iii (2016: 38). She recognises that inappropriate emotional responses can be regarded as irrational, and argues that the real solution to the problem of recalcitrance lies in the fact that we can considerably influence our emotional dispositions (Ibid: 37). Tappolet claims that our emotional systems manifest a high degree of *plasticity* 'in the sense that they are largely shaped, and can also be reshaped, by their socio-cultural environment' (Prinz 2004: 234; Faucher and Tappolet 2008).[57] Therefore, given that our emotional systems have shown that our emotional dispositions are plastic (i.e., they can be modified), it is possible to avoid recalcitrance (via psychotherapy, or by changing our environments, for example).[58]

Tappolet claims that this solution is perfectly consistent with the view that emotions are perceptual experiences (2016: 38). She also argues that this view has the advantage of being general, since all of our emotional dispositions are plastic. However, this aspect of the theory does not stop any of the analogies between emotions and sensory experiences from working.[59] More importantly, according to Tappolet, this difference does not involve the rejection of the claim that emotions can allow us to be aware of values (Ibid: 38). However, in the next subsection I will discuss an argument that aims to reject this claim.

[57] This is acknowledged by biological determinists who argue that basic emotions such as fear and disgust are universally shared as well as innate (e.g., Tooby and Cosmides 1990), as well as by social constructivists who argue that emotions are more complex structures created by socio-cultural groups (e.g., Harré 1986; Armon-Jones 1985).
[58] Concretely, any view of emotions that wishes to resolve the conflict of recalcitrance (as well as other problems resulting from volatile and inappropriate emotions) will benefit from the feature of plasticity of emotions (as seen in Elgin's view).
[59] Nonetheless, it is not clear that the PT (and anti-judgementalist accounts of emotion in general) provides a satisfactory understanding of the undeniable conflict present in the experience of recalcitrant emotions (Helm 2001: 42; Grzankowski 2020: 2). Since my aim in this subsection is only to present objections to the PT, I am not going to discuss alternative theories that claim to provide a better account of the recalcitrance of emotions (e.g., Helm 2001; Brady 2007; Grzankowski 2020).

iii.ii The Evaluative Properties Objection

Brady claims that emotions can lead to a search for reasons that can possibly explain our perceptions (2013: 112-13). For example, the fact that the rollercoaster seems to an observer as really high and fast provides them with at least two reasons to fear it (i.e., that it is really high and really fast). Brady argues that the facts that cause our emotional responses provide the reasons that serve as justification for our evaluative judgements. So, in the previous example, if the observer makes the evaluative judgement 'this rollercoaster is fearsome,' the reasons that would justify both their judgement and their emotional response would be that the rollercoaster is really high and really fast. According to Brady, this explanation contradicts at least in in one sense the analogy that the PT establishes between emotional experiences and sensory experiences. Brady explains that fear on its own does not justify the observer's judgement that the rollercoaster is fearsome, and so emotions do not play the same justificatory role with respect to evaluative judgements, as sensory experiences do with respect to sensory beliefs. For example, says Brady, to see that the football is right before under normal conditions does justify my belief that the football is right before me (Ibid: 113).

Brady argues that even in normal conditions and in the absence of defeaters, emotions do not satisfactorily justify our evaluative judgements (Brady 2013: 86). For example, if in answer to the question 'why do you find Bob lovable?' Mary simply says that she feels love towards Bob, we will find this reply insufficient and probably ask her to provide at least some of the reasons why she loves Bob. Tappolet admits that examples like this one do show a clear difference in justification between emotions and sensory experiences (Tappolet 2016: 39). Nonetheless, she argues that this difference does not entail that emotions do not have any justificatory power at all. It is possible to convince others of what we are feeling by describing the features of what we are evaluating, just as we would point out of the window to convince someone that is raining, at the same time that they are doubting that it is raining (Ibid: 39). Also, according to Tappolet, emotions can play a justificatory role given the fact that the evaluative features that are perceived by emotions supervene on the natural features of the world (Ibid: 39). For example:

> Fearsome dogs usually have sharp teeth and short tempers, for instance. And if a dog has sharp teeth and a short temper, it is likely to be dangerous (at least for a normal human being), and thus fearsome. So, the fact that a dog has these features gives you reason to perceive it as fearsome, for after all, such a dog is likely to be fearsome, that is, to make fear appropriate. Thus, while it is true that emotions differ from sensory experiences with respect to justification, this has more to do with the nature of values[60] than with the nature of emotions (Ibid: 40).

[60] In other words, evaluative properties such as fearsomeness or disgustingness.

As I understand it, according to Tappolet, although there is a sense in which emotional justification is not the same as the justification gained through sensory experiences, she points out that emotional perceptions will be justified insofar as the features that they represent make the emotional perception appropriate. Tappolet also adds that her epistemological claim is modest, given that her claim is that an emotional perception makes the corresponding evaluative judgement *prima facie* justified, which is not a difficult epistemic status to attain (Ibid: 40).

Nonetheless, there is still a second part to Brady's objection. Brady also objects to the claim that emotions are perceptions of evaluative properties, and thus can inform us about such properties. If emotions do not serve as justifications themselves, says Brady, how can it be argued that they tell us something about evaluative properties? Consider Mary's love. How could Mary's love towards Bob inform us about Bob's 'lovable quality,' if Mary's love on its own is meant to justify Bob's loveliness?

Tappolet acknowledges that emotions can misrepresent their objects (2016: 41). The claim of the PT is that emotions only inform us about evaluative properties when they are appropriate (i.e., when they represent things correctly). The question then would be how it could be that an appropriate emotion could inform us about evaluative properties when we have no clue whether or not the emotion is justified. But Tappolet argues that this worry is the same as the worry of whether a belief could be true when we have no idea whether or not it is justified. Moreover, the fact is that most accounts of epistemic justification accept that a belief can be true whilst being unjustified, and justified whilst not being true. Similarly, says Tappolet, 'it may well be the case that an emotion is appropriate even though we have no idea whether or not it is justified. So, emotions can inform us about evaluative properties even though the question as to whether or not they are justified remains open.' (Ibid: 41)

Discussing briefly the problem of recalcitrance, and how the PT deals with it was useful for the purposes of introducing the problem of recalcitrant emotions, which I will also have to deal with in chapters V and VI. Discussing the PT was also useful given that it provides an account of emotion that is paradigmatically epistemic, and so it shows that at least in some circumstances emotions can perform epistemic roles. In contrast, although MES also attributes epistemic roles to typically moral emotions (such as directing our attention to the morally relevant features of a situation, as well as evaluating some situations as either positive or negative), it does not argue that emotions serve as justifications to our evaluative judgements. MES does not make emotional experiences analogous to sensory experiences. According to MES, given their components, emotions can lead to an increase in our understanding of moral evaluative properties (e.g., of rightness), though they are not regarded as perceptual experiences of evaluative properties that would justify our morally evaluative judgements. Rather, MES relies on the fact that emotions can focus our attention, to explain how emotional experiences can *inform* us about the morally relevant features of certain actions/people/situations. It is only in this sense that MES argues that emotions can act as perceptions, but not that they are, in essence, perceptual experiences.

In the following and last section of this chapter, I will discuss one more account that shows another way in which emotions can be epistemically problematic.

iv. Misleading Emotions

Even though emotions can be regulated and educated, there is yet another way in which emotions can distort knowledge and understanding. Contrary to the general claim of the Standard View (i.e., that due to their variability and volatility emotions are cognitively untrustworthy), Peter Goldie has argued that in recent years there has been a popular view that characterises emotions as epistemic assets. Goldie describes this view as excessively optimistic, since, for instance, it tends to ignore that emotions can be *systematically misleading* due to an environmental mismatch (2008: 155). Roughly put, the relevant aspect of Goldie's observation is that emotions are sometimes *impossible to correct*, and this fact presents epistemic problems, such as the disturbance of our epistemic and preferential landscapes (Ibid: 159-160). The problem that Goldie is highlighting is almost the same as the problem of recalcitrance. What distinguishes Goldie's objection is that it points to the fact that although most emotions are plastic, there are some that in certain cases are biologically impossible or extremely difficult to regulate. Before describing these ways in which emotions can mislead us, I will discuss Hume's example of the 'Good Samaritan', as explained by Goldie. The example illustrates the difficulty that lies in correcting our moral sentiments in general. I will then briefly describe Goldie's theory of misleading emotions, and partially agree with his criticism against an excessive optimism towards emotions' epistemic powers. Goldie's target are not moral emotions, but he uses Hume's view to introduce his.

The first thing to notice in the story of Hume's Good Samaritan, is that the moral 'feelings' involved are also referred to as 'sentiments.' Both typical emotions and sentiments focus on specific objects, but emotions are often taken to be occurrent or episodic, and sentiments are dispositional or enduring (Ben-Ze'ev 2010:55). This is not to say that some emotions cannot persist throughout time, and therefore be dispositional (e.g., long-lasting sadness or grief) (Tappolet 2016: 4). It is my view that in such cases, there is no important consequence in referring to certain emotions as sentiments and vice-versa. In other words, there is no difference in labelling for example, *resentment*, either as an emotion or as a sentiment. However, Goldie's view of the ways in which emotions can be misleading concerns emotions only in the *dispositional sense*, given that the emotions which he is discussing appear to permanently belong to our species' psychological mechanisms (2008: 153).

The second thing to bear in mind, is that Hume's Good Samaritan is meant to show how our emotions affect the accuracy of our moral judgements: '… the fact that the Good Samaritan happens to be my son should not influence my moral judgement of his action, even though his kindness and generosity is

more salient for me, and even though I accordingly feel more admiration, just because he is my son (Hume [1739/40], 472). This is not to say that I should necessarily avoid having the feelings[61] proper to my particular relationship with my son, but rather that these feelings should be kept apart from my feelings about the morality of his action.' (2008: 151)

Even if for Hume—as it is widely known— morality involved feeling more than reason, to adequately engage in moral matters required the use of reason and imagination, in order to *correct* our sentiments in at least two ways. Although the kindness and generosity of the Good Samaritan are more salient (i.e., notable or significant) for me, given that he is my son (say), this fact should not entirely influence my moral judgement of his action. I would have to make a distinction between my sentiments towards him, and those towards the morality of his action. If I morally approve of my son's action, the sentiments towards the morality of the action should remain, even if the same action was performed by someone who I have no personal relationship with (or if it was performed by someone who I actively dislike). This first kind of correction then, would concern our particular relations.

The second variation in our moral sentiments to be aware of, is the proximity or remoteness of their object (Ibid: 151). What is nearer to us is more salient or striking than what is farther away, and this can affect our moral judgements. According to Hume, we should correct our moral judgements of a distant situation —or of someone distant— just as we do when we judge that an approaching object does not really get any larger, despite its appearing to be getting larger (Hume [1777], 227-8). This would be the second kind of correction, which involves the contingent proximity or remoteness of the object, person or situation that is morally judged.

Reason plays a role then, in adjusting our moral emotional reactions or sentiments in order to formulate moral judgements without contradictions (Hume [1739/40], 581-2). However, as Hume also argues, correction by reason is usually not successful and our sentiments do not always follow the judgements we arrive at (through reason) (Ibid: 583).

Given this difficulty, we are often motivated to act following our sentiments rather than what reason indicates (i.e., deliberate thinking). Goldie's aim is to show how this phenomenon extends beyond the moral domain, relying on the work of contemporary psychologists[62] which confirms that emotions can both distort our judgements about the environment and about particular situations. In turn, they can also motivate us to act impulsively in a *systematic* way. Certain emotional responses, which are referred to by evolutionary psychology as 'emotion-based heuristics'[63] (Gigerenzer and Selten 2001), used to have specific roles in previous environments. For example, aggression was an appropriate response at a time

[61] Goldie and Hume talk of 'feelings' in this context as equivalent to sentiments.
[62] See Nisbett and Ross 1980; Gigerenzer and Selten 2001; Fessler 2001; Slovic 2007.
[63] The psychologists mentioned in this essay, as well as Goldie, understand emotions as being part of our 'System 1 thinking' (deliberative thinking belongs in turn to 'System 2 thinking'). Goldie refers to 'System 1 thinking' as 'Intuitive thinking' (2008:158).

when humans needed to constantly make sure to physically protect themselves from one another. Goldie points out that the 'rationality wars' consist then, not on whether emotion plays an important role in our reasoning, but on whether we should be optimistic or pessimistic about the epistemic role of emotion-based heuristics in the *environment in which we now live* (2008:155). Goldie argues that there is an 'environmental mismatch' which '*systematically* leads to wrong intuitive thinking, and thus to wrong motives and wrong actions, and furthermore, that this mismatch is *systematically* not easy to detect or to correct through reason, through deliberative thinking.' (Ibid: 155) He goes on to say that this is true of at least three types of emotion: *male aggression*, *fear and mistrust of strangers* (xenophobia), and *male sexual jealousy*. It is easy to see how these emotions used to play significant adaptive roles, whereas today they mainly create social problems (e.g., sexual jealousy can lead to out-of-control violence).

According to Goldie, there is also an *environmental mismatch* when it comes to our moral sentiments. As mentioned earlier in the Humean discussion, the particularity of our personal relationships and how close we are to the object of our moral sentiments—which are part of our intuitive thinking—will give rise to specific epistemic problems. For example, when genocide happens in a country far away from where we live, we typically fail to be properly affected by it (i.e., we are not motivated to do something about it). Psychologist Paul Slovic says: 'System 1 [intuitive] thinking evolved to protect individuals and their small family and community groups from present, visible, immediate dangers. This affective system did not evolve to help us respond to distant, mass murder' (2007: 84). It seems, then, that at least some of our moral sentiments (e.g., empathy and fellowship) will fail to make salient the fact that all lives matter equally, since judging that genocide is morally wrong should motivate us to do something about it and make it stop.

Is it possible to correct systematically misleading emotions (moral or otherwise), such as the ones described by Goldie? Like Hume, psychologists such as Slovic and Gigerenzer have assumed that humans, broadly speaking, have dual processes of thinking (intuitive—which includes emotions—and deliberative which involves reasoning), and our deliberative thinking is meant to regulate the impulsiveness and unstable character of our intuitive thinking. For deliberative thinking to succeed in directing our moral sentiments, the 'good' reasons have to also turn into 'motivating' reasons. It is not enough to feel empathy and come to judge that genocide is morally wrong (given that it terminates with large amounts of meaningful lives). As discussed above, the reasons that are associated with our moral sentiments should also prove to motivate action. Otherwise, deliberative thinking in the moral domain fails: if we really understood how terrible genocide is, then we would try to do something about it (2008:159).

On the other hand, to correct male aggression, sexual jealousy and xenophobia also seems to be a complicated task. The first step towards correcting one of these misleading emotions would be to

recognise them, but this is hard to do given that they skew the *epistemic landscape* (Ibid: 160). In other words, it is not easy to identify them. For example, xenophobia might direct someone's attention to the crimes committed by immigrants (as reported by questionable sources), and ignore reliable data which indicates that the work of immigrants is good for the country's economy. In a case like this, there is little which deliberative thinking can do, since the evidence presented by the emotion has already undermined it without the persons' conscious awareness (Ibid: 160). This person probably thinks that their attitude shows concern for the welfare of their country, instead of an irrational fear of strangers.

The second epistemic difficulty found in trying to correct these emotions, is that they skew the *preferential landscape* (Ibid: 160). This feature refers to how we get motivationally carried away by the emotion. For example, when Mary is in the grip of jealousy even if she became aware of the state that she is in, and even if she were able to recognise the reasons why she should not show violent behaviour, jealousy might still mask the fact that she knows that she should not behave violently, and hence would lead her to act violently regardless.

Goldie acknowledges that it is possible to attempt to correct misleading emotions by stopping and thinking carefully in order to recognise them. Counting to ten, biting our tongue, shifting our focus by consciously trying to turn our attention away from what is triggering us, etc., can be tactics that may prove to be effective every now and then; but sometimes some emotions will have already altered our epistemic and preferential landscapes before we can correct them.

Despite the undeniable difficulty —or even the impossibility— of correcting systematically misleading emotions, it has also been proven that emotions or emotion-based heuristics advance the understanding of our environment. Not all emotions skew our epistemic landscapes, or at least not always. Also, as seen in section ii of this chapter, it is possible to regulate and educate our emotions in different ways in order to avoid emotions from leading us to error and misfiring.

Goldie is right in pointing out that emotions mislead us in complex ways, which in fact appear to be beyond our awareness and control. A moderate scepticism towards the roles that emotions may play in our thinking mechanisms will also allow us to define these roles better, and optimise them and use them to our advantage. In the next chapter, I will discuss a very specific epistemic function that emotions can perform. I will argue that even if emotions can lead us astray, be variable and volatile, recalcitrant, and systematically misleading, they can also direct our attention to our advantage. As mentioned before, in chapter IV my aim will be to argue that certain emotional episodes can direct our attention to the morally relevant features of a given situation, thereby allowing us to increase our moral understanding.

Conclusion

In this chapter, I attempted to defend the following claims: (i) unless they misfire, emotional episodes can at least be partly involved in the process of the production or modification of knowledge and/or understanding, (ii) emotions can be sources of salience (iii) emotions have six components: intentionality, cognition through patterns of salience, personal evaluation, motivation, and feeling, (iv) although emotions sometimes are variable, volatile, and systematically misleading, they can also be refined and regulated.

In the first section of this chapter, I provided a rough description of different theories of emotion. In particular, I described theories of emotion that belonged to the *evaluative* category in order to later locate my view among them. There are evaluative theories which are constitutive (e.g., emotions as judgements), and causal (e.g., emotions as appraisal processes), and some constitutive and some appraisal views are also categorised as *hybrid evaluative-feeling* views. These views regard emotions as complex affective states, and aim to explain the phenomenology of emotions (i.e., the feeling aspect) as well as the distinctive evaluation and motivational power that they can provide. Also, these views can describe emotions as including patterns of salience (i.e., they can highlight certain aspects of a situation and obscure others). Finally, in this section, I described my own hybrid-evaluative theory of emotions. I argued that my view on emotion is very similar to Aaron Ben-Ze'ev's (2010), given that I also understand emotions to be a general and dynamic mode of the mental system that includes various components. I explained that since MES describes several aspects of emotions, it describes them better than other views.

In the second section, I first described three ways in which some authors have argued that emotions can be epistemically helpful: (i) they can be sources of salience (ii) they can help us to evaluate our environment, and (iii) they can motivate to acquire knowledge and understanding. Second, I described the Standard View that in general emotions are variable and volatile and therefore epistemically unhelpful (e.g., they usually lead to irrational behaviour). In response, I agreed with Catherine Elgin's argument that despite their variability and volatility there are several ways in which emotions can be *regulated*. Third, I provided some examples of epistemic emotions (e.g., curiosity, feeling of simplicity) in order to compare them to typically moral emotions. I argued that what makes moral emotions distinctive is that they can constitute, arise from, or be associated with moral judgements (e.g., compassion can arise from the judgement that 'it is right to perform acts of charity').

In the third section, I described the widely held Perceptual Theory of Emotions (PT). According to the PT, emotions are, in essence, perceptual experiences of evaluative properties (Tappolet 2016: 15). Simply put, the PT grounds this claim in an analogy with sensory experiences. Also, according to the PT, the content represented by emotions does not need to be conceptual (e.g., articulated via structured

propositions) (Ibid: 16). Another way in which the PT explains that the content represented by emotions is non-conceptual, is that we do not need to possess evaluative concepts in order to experience emotions (e.g., it is not necessary to possess the concept of the disgusting to experience disgust) (Ibid: 16). According to MES, acquisition of moral understanding via emotional experiences can include both conceptual and non-conceptual as explained by the PT. In subsection iii.i, I discussed the problem of recalcitrance and the way in which the PT deals with it, and in subsection iii.ii, I discussed the evaluative properties objection to the PT. Finally, I distinguished the PT from MES. According to MES, emotions can work as perceptual experiences only in the sense that by focusing our attention they can *inform* us about the morally relevant features of certain actions/people/situations.

Finally, in the fourth section, I described Goldie's view of misleading emotions. According to Goldie, some emotions can be *systematically misleading* due to an environmental mismatch (e.g., xenophobia) (2008: 155). Systematically misleading emotions are extremely difficult to correct, and therefore give us some reason to remain moderately sceptic about the positive epistemic roles that emotions can perform. Although MES focuses on the *positive* roles that certain emotions can play in the acquisition of moral understanding, it does not ignore the ways in which emotions can also be epistemically misleading.

Chapter IV

Attention and Emotion

In this fourth chapter, my aim is to argue that exercising the epistemic faculty[64] of attention is a necessary condition to acquire understanding in general. Given their involuntary, automatic nature, emotions can focus our attention on what we regard as important, but also on what we do not usually or would not otherwise care about. Concretely, in chapter V I will argue that first-hand emotional experiences can focus our attention on the morally relevant features of morally appraised actions, thereby providing the highest level of moral understanding. My view about cases of attention caused by emotional episodes can apply to cases of acquisition of different types of understanding (e.g., political, aesthetic). However, in this thesis I will only focus on moral emotional episodes that, through the adequate direction of attention can both provide the highest level of moral understanding, and also generate a change in moral perspectives (as I will also argue in chapter VI). I will not provide a unified theory of attention, nor argue for a specific view on attention.[65] I will focus instead on some features of attention which show that attention is necessary for understanding. Any view on attention that endorses the claim that attention furthers understanding in general will be compatible with my theory.

In the first section, I will first provide the definition of attention that I will be assuming throughout the rest of the thesis, and briefly discuss how attention relates to consciousness (i.e., self-awareness). Secondly, I will describe four features of attention (selectivity, clarity, phenomenal character, controllability) in order to later focus on the importance that *selectivity* and *clarity* bear on the acquisition of understanding. Thirdly, I will briefly outline how the involuntariness of attention (James 1890) will be relevant for my purpose in showing that it is possible to acquire understanding without making an explicit effort to do so.

In the second section, I will first provide a brief historical account of some positive epistemic roles attributed to attention in order to later argue for a specific positive role that attention can play in the acquisition of understanding in general, but more importantly, of moral understanding. Secondly, I will briefly describe the attentional requirements for the acquisition of understanding, and argue that given the selectivity and clarity that it provides, attention is necessary for the acquisition of understanding. Later in section iv, I will apply these attentional requirements to my view on moral understanding as described in chapter I.

---

[64] By 'epistemic' here, I am referring to attention as a mental faculty relating to the acquisition of knowledge and understanding.

[65] For a general philosophical discussion of attention see Malebranche (1674), Reid (1788), and for a phenomenological discussion of attention see Merlau-Ponty (1945).

In the third section, I will first elucidate the way in which it has been argued that emotions can direct attention (e.g., LeDoux 1996, Brady 2010, 2013), since this mechanism will be crucial in chapters V and VI to describe the roles that some moral emotions can play. Secondly, I will also argue that although desires are mental states that can also direct our attention (Schroeder 2007) and increase our understanding, the fact that emotions *always* involve a feeling component—alongside their involuntary character— makes them better candidates to explain the two instances of acquisition of moral understanding that I have in mind.

Finally, in the fourth section, I will develop in more detail my two main claims regarding emotion and attention: (i) that by directing our attention emotional episodes can allow us to gain understanding, and (ii) that by involuntarily directing our attention to features of a given situation that appear either as new or different, such emotional episodes can lead us to regard them as salient or relevant from there on. In other words, emotions can lead us to *care* about certain things that we did not previously care for.

i. Attention and Consciousness

William James famously described attention as '…the taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, and consciousness are of its essence' (1890: 404). James argued that attention involved disengagement from some things in order to 'deal effectively with others,' and described attention as a mental state that contrasts with confusion and distraction (Ibid). Throughout the rest of the thesis, I will assume that the phenomenon of attention involves all of the aspects that James has included in this description of attention. Concretely, this general definition of attention seems to be compatible with the features of attention that I will regard as most important (i.e., selectivity, clarity, and involuntary selection) in the acquisition of moral understanding.

As further elucidation of James's description of the mental phenomenon of attention, Carolyn Dicey Jennings has pointed out that our usage of attention is centred around the concept of 'the act of mental selection' (2012:536). I follow Jennings when she adds that the term 'mental' implies everything that can be entirely contained within thought and/or memory, and that the term 'selection' implies the 'prioritization of preferred over non-preferred entities.' (Ibid: 536) In section ii.i, I will argue that selectivity and other features of attention will be very important in the process of acquiring any type of understanding (e.g., mathematical, geographical, musical, moral).

Another very important aspect of attention as described by James, is that attention involves consciousness. There is an ongoing debate in philosophy of mind and cognitive science, on whether attention suffices for consciousness in every case (Jennings 2015). I will remain neutral about this

particular issue, given that my view will only concern specific instances where attention *does involve* consciousness. The kind of 'consciousness' that I am interested in, is consciousness as a *self-aware mental state* whereby an agent is both aware of the object that they are perceiving, and aware of the activity of their perceiving the object.[66] Franz Brentano explains this mental act of seemingly double consciousness in the following manner:

> [Every conscious act] includes within it a consciousness of itself. Therefore, every [conscious] act, no matter how simple, has a double object, a primary and a secondary object. The simplest act, for example the act of hearing, has as its primary object the sound, and for its secondary object, itself, the mental phenomenon in which the sound is heard. (1874: 153-4)

The 'object' at play in this thesis, will be certain *morally appraised actions* (e.g., lying to one's friends, helping those in need). In section iii of this chapter, it will become clearer why self-awareness is necessary for my view for acquiring two forms of moral understanding.

I am aware that there might be some cases in which paying attention to certain objects does not involve a conscious act (nor an act of self-awareness). In other words, it is possible to think of cases where understanding acquired through attention can in some cases be subconscious (i.e., not self-aware).[67] Nonetheless, here I will only be interested in conscious (i.e., self-aware) acts of attention brought about by emotional episodes.

For reasons that will become clearer in section iii, it will be important to keep in mind that different emotional episodes involve higher or lower degrees of attention (Fredrickson 1998); and so, each emotion can have a distinct impact on *self-awareness*. For instance, there can be some emotional episodes whereby we can reach a state of self-awareness without attending to any specific object (e.g., Mary can experience guilt, and be aware of her guilt without paying attention to whatever made her feel guilty, and/or without attending to her actual experience of guilt). To use Brentano's example, it is possible to hear (say) the music from the ice-cream van (where the primary object of the conscious act of hearing would be *the music* from the ice-cream van and the secondary object would be *our experience* of hearing the music); however, this conscious act of hearing the music does not seem to entail that we pay attention to it (i.e., we do not need to recognise the tune, talk about it, stop everything we are doing to listen carefully to it, etc., although we would both be aware of it and experience what it is like to hear that tune).[68]

[66] This view on consciousness is also similar to the views of Gennaro (2004), Kriegel (2009), for example.
[67] In section iii of this chapter, I will briefly explain how my view of consciousness relates to phenomenal consciousness.
[68] On the other hand, some authors have argued that attention is neither necessary nor sufficient for consciousness (e.g., Lamme 2003, Block 2007). For a more detailed description on the nature of attention see Watzl (2011).

As mentioned earlier, in section iii of this chapter it will become clear why my view involves self-aware cases of attention. In the following subsection I will briefly describe some widely discussed features of attention, to later emphasise in section ii that attention is necessary for the acquisition of understanding.

i.i Features of Attention

The features of attention that will play at least some role in my view are: selectivity in information processing, clarity, phenomenal character, development of skills, production of behaviour, scope of attention, voluntariness and involuntariness. However, the four features of attention that will be key for my purposes are *selectivity*, *clarity*, *phenomenal character*, and *involuntariness*. These will be the main features required by the two instances of acquisition of moral understanding—via emotional episodes—that I will develop in subsequent chapters V and VI.

Although the psychological literature on the features of attention is extensive (e.g., see Treisman 1964), in this subsection I will briefly describe only some of the features that are widely discussed in the philosophical literature on attention (at least in the philosophy of emotion and in the philosophy of mind), and that are relevant for my view. There are two reasons why I will limit myself to mention only a few features of attention. The first reason is that my thesis fits into a discussion within philosophy of emotion, and the second reason concerns the purposes of my project.

As argued by Christine Tappolet, *selectivity* 'in information processing, be it voluntary or involuntary, is considered to be the essence of attention.' (2016: 34)[69] The selectivity of attention would be exemplified by what James described as focusing on 'one out of what seem several simultaneously possible objects or trains of thought' (1890: 404). Relatedly, it seems that this feature of attention can be exemplified by at least five cases that at the same time involve voluntary or involuntary information processing. For instance, it seems that depending on the *subject-matter*, the objects selected by attention will vary. For example, if a group of botanists goes hiking in the search of new flower species, their attention will very likely pick out the flowers found in the landscape; in other words, their attention would select flowers among all the other objects in the landscape (e.g., plants, trees, animals). This would be a case of voluntary information processing.

Another case of selective attention is tied to *desires* (Schroeder 2007). Put simply, if I desire an ice-cream cone, my attention will be voluntarily directed at the possible ways in which I can satisfy my desire (e.g., among all the street noises, my attention would select the music coming from the ice-cream

[69] See also Duncan (1999).

van). To illustrate how whatever one wants will make certain things salient (and these things in turn will prompt us to act in a certain way), here is Schroeder's example:

> So, when you want a cup of coffee, you find yourself thinking about a wide range of topics, and considerations having to do with these topics strike you in a special, s*alient*, way, and when they do, this is the kind of thing to prompt you to act in a way that is non-alienating (i.e., they don't puzzle you, they seem natural). These are the kinds of thing that are involved in having a *desire* for a cup of coffee, on my account. Moreover, I can say *which* topics your attention will be directed to, which considerations you will find salient, and which actions they will prompt you to do. You will be prompted to do actions which obviously, given your beliefs, promote P, the object of your desire. (2007: 156, emphasis his)

According to Schroeder, to desire is to be in a particular state of mind (or psychological state) for certain reasons (Schroeder 2020, 2007).[70] In section iii of this chapter, I will compare the way in which desires and emotions direct attention, given that I will argue that, due to their features, emotions and not desires allow us to acquire at least two forms of moral understanding. As I will explain in that section, Schroeder's conception of desire is useful to show that there is a certain rational control over our desires, as opposed to the case of the emotional episodes that we experience.

A third case of selective attention would be the one involved, as James suggests, in *trains of thought*. For example, despite craving ice-cream, it is possible that I also constantly think of ways in which I can exercise more, and so my attention can select either the train of thought that involves ice-cream or the train of thought that involves going to the gym and/or going for a run. This and the examples of selective attention discussed above, are all cases that involve voluntary (i.e., deliberate) information processing.

A fourth case of selective attention involving *involuntary* information processing would be the *spontaneous* case. Emotions can easily illustrate this case. As seen in the previous chapter, emotions can be sources of salience (e.g., de Sousa 1987, Elgin 1996, 2008, Ben-Ze'ev 2010) (i.e., emotions can direct our attention to certain aspects of a given situation). Recall as well that each emotion directs our attention to different things (Gibbard 1990). So, for example, if a spider falls on Mary's face all of a sudden, the experience of *fear* can direct her attention to the ways in which she can get rid of the spider. Of course, spontaneous cases of selective attention need not involve emotions. For instance, a running deer in the motorway might cause our attention to focus on avoiding hitting it with the car.

Finally, another case of involuntary information processing of selective attention can also be illustrated by an emotional experience, or more specifically, by experiencing a moral emotion. As described in chapter III, Mary experiences *guilt* after realising that she often treats Bob badly by yelling at him.

[70] Schroeder expands on Scanlon's view theory of Directed-Attention Desires (1998), which explains—roughly—that our desires relate to the particular reasons one has to want something.

Hence, guilt makes salient at least some of the ways in which she has failed Bob as a friend, as well as ways in which she can repair her offence (Gibbard 1990: 126). This is a case of involuntary information processing, given that it does not depend on Mary's decision to experience guilt (i.e., the way in which selective attention takes place is passive, even if Mary is generally disposed to feel guilt).

Notice that both cases of voluntary and involuntary information processing of selective attention can involve two aspects: (i) selection of *information* and (ii) selection of *possible courses of action.*[71] In the spider example, Mary's attention might be selecting the spider's colour and size (i.e., information about the spider), as well as possible ways to respond to the presence of the spider such as throwing the spider out of the window in case it is not venomous, or screaming for help (i.e., certain courses of action).[72]

The other most important feature of attention for my view, is *clarity* (James 1890: 404, Brady 2013:18). For instance, Malebranche, has argued that what keeps our perceptions from being confused and imperfect is attentiveness (1674: 411-12). Authors such as Pillsbury and Wundt have also argued that an increased clarity of ideas is essential to attention (Wundt 1907; Pillsbury 1973). More specifically, the clarity provided by attention will be relevant for my view insofar as it will aid our understanding in general. So, for example, by closely attending to a newly found species of flowers, botanists will understand better different aspects of the flowers (e.g., in which colours they come, in what season they bloom, etc.). Relevantly, psychologists used to understand attention as 'an increased clearness of a particular idea' (Treisman 1964: 12). I will discuss clarity in more detail in section ii.i of this chapter.

Another feature of attention is its *phenomenal character* of our conscious experiencing (Nagel 1974 and Shoemaker 1994). At least, this applies to conscious (i.e., self-aware) acts of attention. In other words, there is something it is like to, for example, concentrate attention on a train of thought, or to play basketball with all of our attention (Watzl 2011: 843). It seems that there is something it is like to focus our attention on the book we are reading instead of, for example, the bee that is buzzing right outside of the window. As Watzl puts it, there is something it is like for us to know what attention is by 'first-person conscious acquaintance, just was we know what, for example, pain is.' (Ibid) The

[71] Wu's theory of attention as *selection for action* (2011) illustrates this aspect of selective attention. According to Wu, agents are usually faced with many inputs, as well as with many possible responses to these inputs (he calls this the *Many-Many problem*). Wu argues that attention should be "identified with the processes involved in solving the Many-Many Problem – namely selection of a specific input to inform a specific response." (Wu 2011: 103).

[72] An example of a theory that focuses on the selective process involved in attention would be Watzl's 'Structuring View' (2010, 2011): 'The structuring view takes as its starting observation the idea that attention is contrastive: it structures our mental life so that some things are in the foreground of others. For example, when someone focuses her attention on a project like writing a book or raising children, she will structure her life so that other things form the background of that project. In order to pin down the functional role of attention as structuring, the idea continues, we must take its phenomenal character seriously.' (Watzl 2011: 849). Other examples include the Selection for Action View (Wu 2011), the Rational Access View (Smithies 2011), and the Cognitive Unison View (Mole 2010). Additionally, these other distinctions have been made amongst different types of attention: Focal vs. Global (Treisman 2006), On-off attention vs. degrees of attention (Watzl 2011), Exogenous vs. endogenous (Smallwood and Schooler 2009), Perceptual vs. executive (or central) (Pashler 1998), and the process of attending to something vs. the event of shifting attention from one thing to another vs. the state the process results in (Watzl 2010 and Wu 2011a).

phenomenology of conscious experiencing will be important to describe the acquisition of understanding via emotional experiences.

Tappolet has also pointed out that there are differences in the *scope of attention*. In other words, attention can zoom in and concentrate on details, or it can zoom out and focus on global features (Tappolet 2016: 34). This feature of attention is worth mentioning given the fact that difference in attentional focus can aid understanding in different ways. For example, the botanists can focus on specific aspects of the new flowers in order to differentiate them from other flowers, but they can also focus on the aspects that make them similar to other plants (e.g., trees).

As suggested earlier in the discussion of voluntary and involuntary information processing, attention can involve *production of behaviour* (James 1890: 447-48). According to James, 'volition is nothing but attention' (Ibid: 447). James accounts for this claim by linking attention to autonomy (Ibid). Whether attention is only an effect of our brain processes, or whether it is at least a cause in part of some brain activity (or both), it is linked to the *effort* involved in attending to whatever it is we want to do (Ibid: 449-454). This feature of attention will also be important in describing how attention caused by emotional episodes can lead to action, and performing certain actions can lead to the acquisition of moral understanding.

Although James' goal when speaking of volition is mainly to explain the nature of the process of attention (as an effect rather than cause), he points out that attention can involve *imagining the things or actions* that one is attending to, or looking for (Mole 2021: 3.4). According to James, the process of attention involves two physiological processes (i) the adjustment of the sensory organs and (ii) the anticipatory internal preparation of the ideational centres concerned with the object to which attention is paid (1890:434). The first process involves physical reactions such as focusing our eyesight on our targets, and the second one accounts for attention's link to imagination.[73] And so, for example, when the botanists go to the mountains to search for the new species of orchids, they are (i) directing their eyesight to the flowers in the surroundings, and (ii) actively looking for the flower that they have pictured in their minds (i.e., imagined); additionally, they might be attentive to the work that they will have to undergo to study and classify the new orchid.

In the next subsection, I will discuss the difference between voluntary and involuntary attention. This distinction will be important for my view, given that *involuntary* attention caused by emotional episodes will explain part of the process of the forms of acquisition of moral understanding that I will elucidate in the next two chapters: first-hand emotional experiences, and change in moral perspectives.

[73] It will be relevant for my view to emphasise the role of imagination in the process of the acquisition of moral understanding, given that it seems to be involved in the process of attending to new information (i.e., morally relevant features of a given situation), and to possible courses of action (i.e., *prima facie* morally right or morally wrong actions).

i.ii Voluntariness vs. Involuntariness

As discussed in the previous section, attention can involve *voluntary* information processing (i.e., active), as well as *involuntary* (i.e., passive) (James 1890: 416). The main distinction between voluntary and involuntary attention, is the fact that voluntary attention always requires that we make an *effort* to attend to an object for the sake of some other interest (Ibid: 420). For example, the botanist who is interested in getting to know new flower species will make an effort to be on the lookout for flowers that are unknown to them. I will also be understanding the voluntariness of attention as its capability of being *controlled* (Watzl 2011: 843). Kant argued, for instance, that the control which we can exercise over our thoughts and our imaginings involves attention (Mole 2021: 3.4). So, for example, when taking a driving test (or any kind of test for that matter) we can exert control over thoughts of going to the cinema by ignoring them.

In particular, the involuntariness of some acts of attention will be one of the most relevant features of attention for the two main claims of this thesis. Authors such as Michael Brady have argued that there is great epistemic value in the fact that attention can be involuntary and *effortless*, all things considered (2013: 19). Brady claims that such automatic shifts of attention in which we are passively drawn to things that are relevant to our concerns, are not epistemically costly as opposed to voluntary attentiveness (i.e., with an automatic attentive response we do not need to actively, consciously, and continually scan the environment to detect things that are relevant to us) (Ibid: 19). Brady also points out that there is another advantage of involuntary shifts of attention which concerns *the speed* of the response. Put simply, involuntary (i.e., automatic) shifts of attention would seem to be quicker than conscious, voluntary, and effortful attentional shifts; moreover, there can be practical advantages in a quick response to, say, potential danger. (Ibid: 19)

Although Brady is right in pointing out that two advantages of involuntary shifts of attention are the fact that they are not cognitively costly, and that they lead to quick responses, there is another advantage about involuntary shifts of attention that will be more relevant for the claims that I will develop in chapters V and VI. The epistemic advantage that I am referring to is that cases of involuntary attention caused by an emotional episode will prove to be useful not in terms of epistemic cost or practical speed, but in terms of *acquisition of understanding that we would otherwise not get*. In the following chapters, I will discuss cases of moral understanding that are not sought out voluntarily.

To sum up, in this section I have briefly elucidated some features of attention that will play at least some part in my view, as it will become clearer in this and in the next chapters. These features are: selectivity in information processing, clarity, phenomenal character, difference in scope, development of skills, production of behaviour, voluntariness and involuntariness. In the following section, I will provide a brief historical account of epistemic functions attributed to attention in order to support my

claim that attention is a necessary condition for the acquisition of understanding. This claim will be very important for my view, given that I will argue that since emotions can direct our attention, and attention is a necessary condition for the acquisition of understanding, emotions can promote the acquisition of understanding at least in some cases.

ii. Brief Historical Overview of the Epistemic Functions Attributed to Attention

By briefly recounting some ways in which, historically, various authors have attributed certain epistemic roles to attention, my aim is to provide additional reasons to endorse the view that attention is a necessary condition for the acquisition of understanding.

According to Descartes' *Third Meditation*, only when *attention* is being paid to clear and distinct ideas (i.e., *prima facie* true ideas), is it possible to stop doubting them (1641: 'Replies to Objections', 309). In other words, in Descartes' epistemology, the transition from the state of radical doubt to the state of certainty about the truth of clear and distinct ideas is explained by the intervention of attention.

Relatedly, in the second book of his *Essay Concerning Human Understanding*, Locke categorised attention as a 'mode of thinking' (together with contemplation, study, remembrance, dreaming, etc.) (1698: II, 19 §1). Locke's treatment of attention as simply another way in which our mind works, suggested that an independent theory of the faculty of attention is not needed once a theory of thinking has been established. Put simply, Locke presented the phenomenon of attention as acting upon already memorised ideas.

However, during the eighteenth and nineteenth centuries some authors began to define attention differently. For instance, according to Henry Home Kames, attention is involved in the acquisition of ideas:

> 'Attention is that state of mind which prepares one to receive impressions. According to the degree of attention objects make a strong or weak impression. Attention is requisite even to the simple act of seeing.' (1769, 18)

On the other hand, Dugald Stewart argued that attention is also involved in the production of at least some skilled behaviours (e.g., juggling requires the juggler to attend properly to various objects in motion) (1792: 53, 62). Still, Stewart also endorsed Locke's claim that attention plays a role in recalling memorised ideas. Nonetheless, attention started to be seen as a more complex epistemic faculty, and which played a role in the explanation of perception, skilled action, and memory.

By the end of the nineteenth century, some psychologists continued to investigate the epistemic functions of attention. For instance, G.F. Stout argued that theories of attention needed to explain the function of attention in reflective thought (1891), and Alexander Bain argued that the essential feature of attention was its influence on action (1888).

William James also described attention as a process that involves 'adjustment of the sensory organs' (e.g., focusing our eyes on a target), and imagining the things or actions that one is attending to (e.g., when playing an instrument, members of a band need to imagine the singer's part before playing the rest of the melody, thereby closely attending to the progression of the music) (1890: 411). Nonetheless, authors like F.H. Bradley still thought, like Locke, that since attention is a thinking mode like any other, there really are no specific attention-processes to be identified (1886: 316).

Later, when behaviourism became popular at the start of the twentieth century, the study of attention became neglected (Mole 2021: 1.6). Later on, cognitive psychology emerged, and throughout discussions and discoveries regarding the limited capacity of the brain to process information and respond to different stimuli, the role of attention seemed relevant once again (e.g., see Broadbent 1982). According to Donald Broadbent, there seems to be a 'bottleneck' in our information processing capacity, and so the role of attention, roughly put, is to control this bottleneck by shifting our focus from one source of stimulus to another (1982: 253). Broadbent's view on attention influenced later theories of attention developed in the twenty-first century, especially those under the category of *Capacity-Limitation Theories* (Mole 2021: 2.1).[74]

Attention's apparent role in preparing our mind for the acquisition of ideas, providing clarity of ideas, processing information, and production of skilled behaviour—as described by some of the authors above—provides at least some reason to endorse the claim that attention is a necessary condition for understanding. In the next subsection, I will argue in more detail that given some of its features, attention seems to be required for the acquisition of understanding.

ii.i Attentional Requirements for the Acquisition of Understanding

Before explaining the sense in which attention is necessary for the acquisition of understanding, it is important to recall that in chapter I, I have endorsed a reductionist account of understanding-why *p*. In other words, understanding-why *p* is no different from knowledge-why *p*. More precisely:

[74] As Brady notes (2013: 16), the psychological literature on attention is extensive. See for example: Pashler (1998); Johnson and Proctor (2004).

> An agent understands why *p* if and only if she has a sufficient amount of knowledge why *p* (Sliwa 2017: 530).

Knowing why *p* can involve deferring to the testimony of others, reasoning whereby one is able to follow explanations why *p*, being able to draw distinctions and conclusions about *p*, performing certain actions, etc.

Overall, Mary's capacity for moral reasoning consisted of five abilities: (i) being able to distinguish morally right actions from morally wrong actions in pictures both with descriptions and without descriptions (ii) providing autonomous explanations as to why she considered something to be right or wrong (iii) drawing distinctions among different kinds of right actions and among wrong actions (iv) giving various examples of right and wrong actions, and (v) imagining or providing tentative answers to novel moral dilemmas.

Ultimately, I will defend this kind of reductionism[75] to argue that instances of moral understanding why *p* seem to be acquired in various ways, given that moral knowledge can be acquired via different cognitive processes such as imagination, intuition, affective responses, moral reasoning, etc. (Ibid: 548) So, for example, when Mary experiences guilt for the first time, she may acquire new understanding why lying to Bob is wrong, given that she already has some knowledge why lying is wrong, and given that she already knew that lying is *prima facie* wrong (through a reliable cognitive process, presumably). Moreover, reductionists are committed to the claim that every instance of understanding is constituted by some degree of knowledge (Ibid: 530). In other words, every possible instance of understanding-that, understanding-how, understanding-why, etc., is reducible to a sufficient amount of knowledge. Henceforth, by 'understanding' I will be referring to the different instances of understanding that as human beings we are capable of acquiring.

There are at least two ways in which it can be argued that attention is required for the acquisition of understanding. The first one involves attentional requirements for *perception* (Mole 2021: 3.4).[76] For instance, 'Inattentional blindness' experiments (Mack and Rock, 1998, 2003), in which participants are visually presented with a picture that gradually changes whilst they are asked to attend to a specific part of the picture (e.g., to check whether the vertical lines of a cross are longer than its horizontal lines) show—among other things— that inattentive observers can fail to notice perceivable objects that attentive observers would find obvious. The experiment also shows that observers under normal conditions can fail to perceive clearly noticeable objects when these objects are unexpected. Put simply,

[75] The link between knowledge and attention that I will be focusing on, is the link between attention and *acquisition of understanding*. I will not be discussing the apparent links between attention and demonstrative reference (see Campbell 2002), and attention and knowledge of other minds (see Moore and Dunham 1995).
[76] Recall that Kames argued that 'attention is requisite even to the simple act of seeing' (1769: 18).

the experiments serve as good evidence that acquisition of perceptual understanding presupposes attention.[77]

The second way concerns *a priori* reasoning. Similarly, there seem to be 'attentional demands' that have to be met at least in some occasions before an agent's grasp of a thought is regarded by the agent as a true belief (e.g., in order for Mary to think that making jokes about Bob is wrong, she needs to be attentive to reasons that justify this belief, such as that it would make Bob sad) (Mole 2021: 3.4). It is important to note that these attentional requirements are not merely for alertness' sake:

> 'They are not captured merely by saying that, in order to gain knowledge, the thinker has to pay some attention to the relevant ideas. A thinker may be attending to a syllogism, but, if he is attending to its rhythm, he may still be unable to see that the conclusion follows.' (Ibid: 3.4)

Nonetheless, there still appears to be no general epistemological theory for the attentional requirements for the acquisition of understanding. However, the cases discussed above (i.e., attention required for accurate perception and attention in *a priori* reasoning), alongside the historically attributed epistemic functions of attention can support to a great extent the claim that attention is necessary for the acquisition of understanding. Moreover, at least two features of attention discussed in section i.i are useful to strengthen this claim:

(i) Selectivity. As discussed in section i.i, selection is considered to be the essence of attention (Tappolet 2016). Attention can be selective in different ways (e.g., in voluntary and involuntary information processing, it can be caused by desire and by emotion, etc.). Concretely, the selectivity of attention is best explained by James as the focusing on 'one out of what seem several simultaneously possible objects or trains of thought' (1890: 404). As shown by the inattentional blindness experiments, without an observer's attentive *focus* clearly perceptible objects can go unnoticed. This suggests that to perceive any given object or situation, our attention needs to be selective and focus on the object or situation at stake. Without such focus, it would be impossible even to perceive whatever we are trying to understand. Hence, the selectivity of attention seems to be a necessary condition for the acquisition of understanding in general. Recall the botanists' example. If a botanist is looking for a rare orchid in the mountains in order to study it and learn more about it, their attention will likely select the flowers that look similar to orchids in general. (i.i.) Selectivity in the production of behaviour. Following James (1890: 434), it was also discussed that in some cases of voluntary and involuntary information processing, attention can involve the activity of imagining a set of possible actions. Given that, as seen in chapter I, the acquisition of understanding—among other things— can involve ability knowledge

[77] See also Neisser 1979; Most et al 2005.

(i.e., knowledge-how), selecting certain courses of action can increase (or decrease) understanding of a given matter. For example, botanists know how to do different experiments in order to discover the interactions that plants can have with the environment (e.g., by plucking them out of their original soil and then planting them somewhere different); in turn, the results of such experiments can improve the botanists' understanding of the general plant's behaviour.

(ii) Clarity. Recall that attention provides an increased clarity of ideas (Wundt 1907; Pillsbury 1973), and prevents us from confusion (Malebranche 1674). James has also explained that the clarity produced by attention is a *condition* for intellectual discrimination that involves comparison, memory, and thinking of relations between objects (i.e., internal analysis) (1890:426-27). Clarity seems necessary for distinguishing one idea from another (e.g., the idea of an orchid from the idea of a poppy), and therefore without discriminating and comparing different flowers, the agent cannot be said to understand them (e.g., to understand why orchids grow under certain conditions).

Paying attention is presented by James as a condition for the acquisition of skills. James argues that children should be instructed '…in such a way as to knit each new thing on to some acquisition already there; and if possible, awaken curiosity, so that the new thing shall seem to come as an answer, or part of an answer, to a question pre-existing in his mind.' According to James, the longer one attends to a topic the more 'mastery of it one has' (Ibid: 424). Although James does not provide an account of the type of skills that are acquired by continuously attending to any given topic or by being able to concentrate, he argues that the immediate effects[78] of attention make us: perceive, conceive, distinguish, remember, and react quickly (shortens 'reaction-time') (Ibid: 424); all of which will arguably can allow us to acquire understanding of whatever we choose to attend to (Ibid: 424).

Now that I have described the ways in which attention is necessary for understanding, in the next subsection I will begin to elucidate in which way can emotions direct our attention.

iii. How Can Emotions Direct Attention?

In chapter III, I argued that different emotions direct our attention to different things. Roughly put, I pointed out that emotions can involve patterns of salience (i.e., each emotion makes specific objects seem noteworthy) (de Sousa 1987, Gibbard 1990, Elgin 2008, Ben-Ze'ev 2010, Tappolet 2016). For

[78] James points out that attention's 'remote' effects are too incalculable to be recorded (1890: 424).

example, *guilt* directs Mary's attention to the actions whereby she wronged Bob, and *fear* directs Mary's attention to the size of the spider, as well as to possible ways in which she can get rid of it. Some authors have argued that this feature of emotions is epistemically valuable in the sense that they allow us to perceive significant changes in our situation (e.g., the sudden presence of venomous spiders) (Ben-Ze'ev 2010: 42), make us aware of features of a situation that we would otherwise overlook (e.g., fear of venomous spiders can prompt Mary to notice whether there are any in sight), and even provide insight into these features (e.g., fear of spiders can make Mary aware of their movements and predict their behaviour) (Elgin 2008: 44).[79] I also argued that given the fact that emotions can play epistemic roles (e.g., they can provide epistemic access to certain properties that we would otherwise overlook), they can advance our understanding.

In this section, I will briefly describe two ways in which emotions can direct attention, following Ronald de Sousa (1987) and Michael Brady (2010, 2013). The theories of both authors are compatible with my view of emotions as mental states that involve an evaluation (Ben-Ze'ev 2010). Later in subsection iii.i, I will discuss the importance of the different ways in which each emotion directs attention, following Tappolet (2016) and Fredrickson et al (1998). This distinct aspect of emotions will have important consequences for my view, since apparently not all emotions would be useful for acquiring moral understanding (e.g., anger can direct our attention to hurting the target of our anger, and this is inappropriate).

I will begin by describing Ronald de Sousa's view. He has provided an explanation as to how emotions can direct attention (1987: 116-121):

> Emotions having targets typically involve a *focus* of attention, which is the apprehension of some (real or illusory) *focal property* of the target. Under certain conditions, which define the standard case, the focal property is also the *motivating aspect* of these emotions. (1987: 117)

The 'motivating aspect' that serves as focal property—as described by de Sousa—refers to an *attribute* of the target, which the emotion identifies as salient. For example, a guitar (target) out of tone (attribute) might be the focus of Mary's emotional response of *contempt.* However, according to de Sousa, some emotions like contempt may or may not have a proper target, and may or may not focus on a real (i.e., non-illusory) attribute. It could be the case that, experiencing contempt, Mary is wrongly focusing on the guitar's tune, given that the attribute of the target that is actually causing her contempt is not the tune, but the person who is playing it, who she does not like; it might be that the guitar is not even out of tune. In a case like this, the person playing the guitar, and not the tune, would be the cause of Mary's contempt. So, although the tune coming from the guitar is the focus of her contempt, this is disqualified

[79] The work of psychologists Derryberry and Tucker (1994) supports these claims. They have argued that emotions "serve to regulate orienting, directing attention toward perceptual information that is important or relevant to the current state" (1994: 170). See also Damasio (1994: 197).

from being the attribute (i.e., motivating aspect) that is really causing her contempt. This illustrates another way in which emotions can be said to epistemically misfire, by directing our attention to the wrong features of a situation. The case I just described would be an epistemic failure, in the sense that Mary would form the false belief that the guitar is out of tune. Nonetheless, the relevant aspect of de Sousa's view for my purposes, is that it shows how it can be the case that emotions can focus our attention on certain targets, given their attributes (whether they are illusory or non-illusory). Whereas de Sousa's view describes —among many other things—a way in which emotions focus our attention on certain targets, it does not involve the description of a process whereby understanding is acquired through emotion.

On the other hand, Michael Brady's view is closer to mine, in the sense that it focuses on how—at least in some cases—emotions can allow us to acquire understanding of the objects, situations, actions, etc., that we evaluate. Brady has also described emotions not only as directing or capturing our attention, but also as *consuming* it (2007, 2010, 2013). Brady has argued that the epistemic role of emotions consists in promoting the understanding of our evaluative judgements, by capturing and consuming our attention (Brady, 2013:147). Roughly put, according to Brady's view, acquiring understanding of a given situation through emotion involves a search for *reasons* that enable us to assess our initial judgement of it more accurately (2013: 129). For example, through the experience of fear, there is usually an appraisal of some threat. In cases where fear does not trigger an immediate—or as it is often said, instinctive— reaction, fear seems to fix our attention on the threat which we would initially, for example, judge as 'dangerous'. According to Brady, this consumption of attention generated by fear can involve a search for reasons that justify some object or event as being dangerous. These reasons would be considerations which allow us to understand the seeming dangerousness that was initially appraised by fear (e.g., is the spider venomous? Is the spider big enough to hurt me?) (Ibid: 129).

Therefore, upon reflection (i.e., search for reasons) caused by fear, Mary can reassess her initial judgement 'spiders are very dangerous.'

Brady describes this search for reasons that emotion promotes, as a route to understanding or gaining insight into our evaluative judgments. He defines emotional experiences as 'proxy' or 'pro tempore' reasons for evaluative judgements (Ibid: 130). He says that when we form beliefs on the basis of emotional experience '…we are believing on the basis of *pro tempore* reasons, reasons "for the time being", which we rely upon precisely because we presently lack awareness of features that constitute genuine reasons for our judgements.' (Ibid: 130)

He suggests our emotional experiences do not serve as 'genuine reasons' (e.g., Mary's initial fear of spiders) given that, in normal conditions, emotional experiences do not represent sufficient reasons for evaluative judgements; it seems that they act as substitutes for reasons that would *justify* entirely judging why some object or event has some evaluative property (e.g., Mary's fear is a proxy reason that justifies

her judgement 'spiders are very dangerous'). Brady's 'genuine reasons' would constitute those reasons which our emotional responses are supposed to be sensitive to all things considered. For example, my experience of shame for having told jokes at a funeral serves as a proxy reason for justifying my judgement 'I have done something shameful;' at the same time, my shame consumes my attention on my deeds at the funeral, and it can make me aware of the fact that a funeral is a solemn occasion, and this is a good reason—a genuine reason— to confirm that my telling jokes was shameful.

Take another example of Mary's experience of *guilt* after having insulted someone at a party. As seen in chapter III, the object of guilt is the offense committed against somebody else (and leads to motive the repair of the perceived transgression). Guilt is also typically associated with the evaluative judgement 'I have wronged someone else.' Following Brady's view, guilt would tend to consume her attention, and lead her to search for reasons that test the accuracy of her judgement. Features of the facts, such as having been pushed down the stairs and being yelled at, might then serve as reasons that warrant or disprove her belief that she acted *genuinely wrong*. After she reassesses her initial judgement in this way, she might change it to 'I have not wronged someone else.'

According to Brady, the role of attention in emotional experience is a lot more persistent than in sensory perceptions, and thus facilitates consciously reflection about emotional objects or events (Ibid). On the other hand, as seen in chapter III, it is sometimes argued that emotions *undermine* reflection rather than motivate it. To support his claim, Brady relies on empirical evidence provided by neuroscientist Joseph E. LeDoux, which indicates that emotions involve *cortical arousal* (LeDoux, 1996). LeDoux describes in his book *The Emotional Brain*, that 'arousal locks you into whatever *emotional state* you are in when arousal occurs.' (1996: 290, emphasis mine). Following LeDoux, Brady argues that persistent cortical arousal generated by a given emotion brings about attentional focus, which in turn facilitates a better evaluation of our emotional situation through promoting conscious awareness of, and reflection about, our emotional circumstances (2010: 121, 2013: 105-6).

Although Brady's view shows nicely that emotions can play a role in the acquisition of understanding by promoting reflection of our evaluative judgements—through the consumption of attention— it only shows one way in which emotions can promote understanding: they can lead us to search for reasons that can serve as justification of our evaluative judgements, and this process allows us to understand better these evaluative judgements, given that we reassess them. In section iv, I will describe other ways in which emotions can provide understanding by directing and focusing our attention.

There is also an important complication for views that argue that emotions perform a certain positive epistemic role through the direction of attention. As Tappolet notes, there is a huge variety of emotions, and the relation between emotion and attention shows, for instance, that what appears to be true of one kind of emotion is not necessarily true of others (2016: 5). For example, *fear* focuses our attention on what we are afraid of, but *boredom* causes our attention to drift away (Ibid: 5). Experiments in cognitive

science have suggested that positive emotions such as *joy* widen our attentional focus (Fredrickson et al 1998: 307). In other words, positive emotions *broaden* the range of thought and action tendencies (e.g., joy can lead someone to perform generous acts, and to think of different ways to perpetuate feeling joyful) (Fredrickson and Branigan 2005: 3). On the other hand, negative emotions such as *anxiety* generally narrow our attentional focus (Ibid: 3), calling forth specific action tendencies and thoughts (e.g., avoidance and thoughts of worry) (Ibid: 3). However, it is not clear that all positive emotions always expand our attentional focus, and that all negative emotions always narrow it. For example, in the case of *interest* (i.e., a positive emotion), our attention can be oriented towards a specific object, and be maintained by this interest (in botany, say). Also, in the case of *disgust* (i.e., a negative emotion), our attention often seems to shift away from the object of our disgust (e.g., mouldy bread). This complication will have an impact on my view on moral emotions, but I will deal with it in chapter VI.

In the next subsection, I discuss the two ways in which emotional episodes differ from desires, in order to argue in chapters V and VI that certain emotional episodes—unlike other mental states such as desires—can lead to two forms of acquisition of moral understanding. I will go back to compare Brady's view with mine in section iv.

iii.i Emotional Episodes vs. Desires

In the following subsections, I will briefly describe the nature of emotions and desires in order to compare them. In section i.i, I made reference to Schroeder's example of desiring a cup of coffee (2007: 156). Desire for a cup of coffee will likely *direct our attention* to considerations that will bring about the satisfaction of this desire. For example, early in the morning such a desire will probably focus our attention on the coffee shops that are available on our way to work. Theories of attention-based desires (e.g., Scanlon 1998, Schroeder 2007), might lead us to assume that—given that they are similar mental states— emotions and desires are equally efficient in directing or focusing our attention in a way that can lead us to acquire understanding (by making relevant information salient, for instance).

To avoid this implication, I will argue in section iv that the *feeling* component of emotions renders them epistemically privileged in the sense that they seem to focus our attention in a distinctive way. But first, in subsection iii.ii, I briefly mention again the components of emotion that I have already described in chapter III, in order to elucidate the way in which they can direct our attention by making us interested in certain actions, objects, environmental and psychological features (ours and those of others). I describe the nature of desires and compare it to the nature of emotions in subsection iii.iii, in order to show that desires are generally less efficient in making us interested in different aspects of different situations, due to their lack of a necessary feeling component.

iii.ii Emotional Episodes

In chapter III, I argued that my understanding of emotions was most compatible with Aaron Ben-Ze'ev's general definition of emotion as a *mental mode*:

> An emotion is a general mode (or style) of the mental system. A general mental mode includes various mental elements and expresses a dynamic functioning arrangement of the mental system (…) This mode involves cognition, evaluation, motivation, and feeling. (2010:57)

Following Ben-Ze'ev, I also elucidated six components of emotional episodes: *intentionality* (e.g., Mary's anger is directed *at* Bob), *cognition through patterns of salience* (e.g., Mary's anger directs her attention to the ways in which Bob is acting strange, and this involves her understanding Bob's behaviour as mildly or strongly offensive), *personal evaluation* (e.g., firstly, by experiencing anger at Bob, Mary assesses Bob's actions as negative, and secondly, she clearly has a personal concern for Bob, and that is why she is not indifferent to his behaviour), *motivation* (e.g., anger might dispose Mary to act in an antagonistic manner by avoiding him or challenging him), *feeling* (e.g., Mary's anger at Bob might involve a headache or a particular state of uneasiness), and *dynamism* (e.g., Mary's experience of anger might have been caused by a memory and/or a thought of having been wronged by Bob, or her anger can itself be the cause that leads her to evaluate Bob's action as wrong). In section iv of this chapter, I will argue that given these components, emotional episodes can be indicators of what we care about, and they can be a source of information of a given situation.

According to Ben-Ze'ev, emotions are *occurrent* mental states as opposed to *dispositional* (2010: 43). Emotional episodes typically occur when we perceive positive or negative significant changes in our personal situation, or in the situation of those related to us. Other occurrent mental states include *moods*, but unlike emotions moods lack intentionality[80] (2010: 55). Dispositional mental states include enduring sentiments with specific intentionality (e.g., love or grief for someone), and affective traits with general intentionality (e.g., shyness) (Ibid: 55). Importantly, other intentional states such as beliefs and *desires* can be dispositional[81] in the sense that, even if I do not attend to this belief or desire, I can still be described as having them (Ibid: 55).

This characterisation of desires is compatible with Michael Smith's Humean dispositional conception of desire (1987), which I will elucidate further in the next subsection.

[80] See Kriegel (2019) for discussion.
[81] Desires can also be occurrent and intentional (e.g., at the moment, Nora desires *tea*) (Schroeder 2020: 2.4). According to Schroeder, dispositional or standing desires are desires that 'one has that are not playing any role in one's psyche at the moment. Occurrent desires, on the other hand, are desires that *are* playing some role in one's psyche at the moment.' (Ibid: 2.4)

iii.iii Desires

My aim in this subsection is twofold: (i) to briefly describe current thought on the nature of desires, and (ii) to argue that although desires and emotions are similar kinds of mental states, and although both can serve as psychological mechanisms to focus our attention and gain understanding, an important difference between desires and emotions is that emotional episodes *always* occur with a feeling component and this makes them more reliable directors of attention.

Desires are mental states that are commonly associated with acting, feeling, and thinking in certain ways (Schroeder 2020: intro). For example, if Mary desires a cup of coffee, she will probably make herself a cup of coffee; if she does not make herself a cup of coffee straight away, she will probably feel the urge to do so. Also, she will probably find herself often thinking about a hot cup of coffee. Desires are usually understood as a broad category of mental states, or more specifically, as *pro attitudes* (Smith 1987: 55).

Smith has described desires as mental states which have phenomenological and propositional content:

> For ascriptions of desires, unlike ascriptions of sensations, may be given in the form 'A desires that p', where 'p' is a sentence. Thus, whereas A's desire to may be ascribed to A in the form A desires that he *s*, A's pain cannot be ascribed to A in the form A pains that p. (1987: 47)

Desires have phenomenological content just to the extent that the experience of certain feelings is one of the things that they are disposed to produce under certain conditions (Ibid: 55).

However, long-term desires do not involve phenomenological content and we still regard them as desires (Ibid: 49). Mary can have the desire to learn more about moral theories, and sometimes feel excitement or curiosity associated with her desire, but there might be other times when she is not feeling excited nor curious whilst still having the desire to learn more about moral theories and talk to experts in the subject or buy books on the topic. This example illustrates the fact that desires may be had without being felt (Ibid: 49), and that desires can be *dispositions* to act in certain ways (Ibid: 53).[82] Roughly put, Smith's view of desires describes them as a broad category of psychological states which are dispositions to act in certain ways under certain conditions (i.e., they have functional roles) (Ibid: 55). This view of desires is compatible with Ben-Ze'ev's, and which I am endorsing.

Here I will start to point out the relevant differences between emotions and desires for my purposes. From what has been argued so far in section iii, we can identify three: (i) emotional episodes are

[82] Smith argues against views that describe desires as essentially phenomenological states (e.g., Platts 1979), mainly by pointing out that sensations or feelings have no propositional content (1987: 49).

occurrent mental states and desires can be both occurrent and dispositional mental states, (ii) emotional episodes—as I am understanding them here— have six components whereas desires are realizable in multiple ways, and (iii) emotional episodes necessarily have a feeling component—what Smith has termed 'phenomenological content'—whereas desires only have it under certain circumstances.

Following the characterisation of emotion given in the previous subsection, it can be argued that desires also involve intentionality (e.g., Mary desires that Bob teaches her how to dance) (Strawson 1994); personal evaluation (e.g., Mary's desire that Bob teaches her how to dance regards dancing as something good or worth-having) (Stampe 1987; Scanlon 1998; Oddie 2005; Schroeder 2007); motivation (e.g., Mary's desire disposes her to take actions that are likely to bring about Bob teaching her how to dance) (Anscombe 2000; Smith 1987); and dynamism (e.g., Mary's desire can generate different actions such as watching Bob dance, but watching Bob dance in the first place can also trigger her desire that he teaches her how to dance).

It can also be argued that, like emotions, desires can involve cognition. Ben-Ze'ev describes the cognitive component of emotion in this way:

> The cognitive component supplies the required information about a given situation. No emotional attitude toward something can emerge without some information about it. The cognitive component in emotions is often distorted. This is due to several related features typical of emotions: (a) partiality, (b) closeness, and (c) an intense feeling dimension. (2010: 47)

Thus, when Mary experiences *guilt* for having yelled at Bob, her guilt makes salient certain facts about Bob's personality: that he is generally absent minded, that he is usually kind, and that he has always been a generous friend. This information about Bob (towards whom Mary is experiencing guilt) is brought about by Mary's episode of guilt. Nonetheless, as Ben-Ze'ev points out, this information supplied by Mary's guilt about Bob's personality might be distorted. Mary's information about Bob's character might not match the results of Bob's psychometric test, given that she is already partial to Bob, being her friend. Also, her close relationship to Bob might not make her notice that Bob is not kind with the people he works with, and also, the feeling that comes with her guilt might be too intense and override her ability to assess Bob's character correctly (Ibid: 48).

Similarly, an account of attention-based desires can describe desires as involving a cognitive component. One of the effects that desires have on attention is directing our attention to relevant information. So, for example, if Mary desires to learn about moral theories, then her desire will direct her attention to opportunities to gain information about them, to people who discuss them, and so on (Schroeder 2020:1.4). Desires can also involve cognitive distortion in the same way as emotions (i.e., given partiality, closeness, and intensity of feeling). However, since one of the features of desire is that it does not necessitate a feeling component, this would appear to make desires less problematic in terms

of information acquisition. Nevertheless, below in section iv I will emphasise the importance of the feeling component of emotion in the direction of attention, and therefore in information and the acquisition of understanding.

There is a final and important distinction between desires and emotions, which involves *rational control*. Simply put, given the automatic character of emotion (Ben-Ze'ev 2010; Brady 2013; Tappolet 2016), there is no—at least— immediate control to be exerted over the emotional episodes that we experience. It seems odd to suppose that Mary can *choose* to experience an episode of guilt, when the fact is that she was overcome with it given certain circumstances. In contrast, desires can dispose agents to act upon them, depending on their strength (Schroeder 2020: 2.3); as Schroeder argues,

> Desire strength could be determined by the amount of pleasure or displeasure apparent satisfaction of the desire would bring, or by the degree to which a state of affairs seems good, or by the degree to which one's attention is drawn to the reasons to bring some state of affairs about, or by the amount of reward-based learning apparent satisfaction of the desire would cause, or by some average of all of these. (Ibid: 2.3)

Therefore, one can find different reasons to act based upon a desire and have at least certain control on deciding which actions to perform. On the other hand, the strength of emotional episodes is determined by the intensity of their feeling component, and reasons or consideration to perform certain actions can be overridden given the intensity of feeling (Ben-Ze'ev 2010: 48).

Although there are also *occurrent* desires (as opposed to standing and dispositional), which can be made analogous to [automatic] emotional episodes, they do not necessarily override other considerations. (Schroeder 2020: 2.4). For example, my desire to have another ice-cream can be occurrent even as I decide to walk away from the ice-cream van. Hence, there are ways in which desires can be controlled and emotional episodes cannot.

To summarise, emotions and desires differ in at least two ways: (i) emotions can be described as a more complex mental state (with six components) whereas desires appear to be a broader mental category), and (ii) unlike desires emotional episodes always involve a feeling component.

In the next and last section of this chapter, I will argue that the feeling component of emotions (together with its other components) renders them as distinctively useful—if not the best—mental state to focus our attention, and therefore can allow us to gain understanding of the object, action, or state of affairs that they are reacting to.

iv. Emotions, Attention, and Understanding

In the first sections of this chapter, I argued that attention is necessary to acquire understanding. In this section I will argue that, although emotions *are not necessary* to acquire understanding, given their components, emotions can focus our attention in a way that can allow us to gain understanding. To do this, I will first describe the kinds of understanding that can *prima facie* be acquired through certain emotional episodes given the six components of emotions. Secondly, as mentioned in the introduction to this chapter, I am also interested in describing the way in which emotional episodes can make us consciously *care*—and therefore become aware— about certain aspects of a given situation that did not seem relevant or salient to us. I explain this phenomenon by turning to the evaluative and feeling components of emotion. Hence, in subsection iv.ii, I describe how an emotional episode typically presupposes an interest on whatever it is that agents react to (Reid 1969: 76-77), and explain the relevance of the feeling component for the cases of acquisition of moral understanding that I will elucidate in the next two chapters.

iv.i Emotions and Understanding

In chapter III I discussed some ways in which some authors have argued that emotions can be epistemic assets. For instance, according to Elgin— at least when reliably correlated with the objects they represent— emotions advance our understanding providing information about the environment (e.g., fear is an indication of danger) (2008: 35). According to Johnston and McDowell, emotional responses can provide epistemic access to certain kind of properties e.g., the admirable, the contemptible and so on (Smith, Lewis, and Johnston 1989, McDowell 1985). According to Tappolet emotions are perceptual experiences of evaluative properties (e.g., unless it misfires, disgust would consist in perceiving something as disgusting) (2016: 15). Finally, some examples of putative epistemic emotions are curiosity, surprise, joy of verification, and disappointment or joy of falsification (Scheffler 1977).

The claim that emotions provide information about the environment is compatible with my view, given that it is explained by the *cognition* component of emotions. However, as I mentioned in chapter III, the crucial positive epistemic feature of emotions for my view is that emotions *focus our attention* on the aspects that—depending on the emotion—are seemingly worthy of notice in a given situation (e.g., empathy focuses our attention on the suffering of our friend) (2008: 43). I have described this feature of emotions as epistemically useful, in the sense that direction of attention can lead to the acquisition of understanding. In this chapter, I have argued that attention is necessary for the acquisition of understanding (given its selectivity and the clarity that it provides), to support the claim that since

emotions can direct attention, emotions can lead to understanding. So far, I have only explained in more detail how can this be the case according to Brady's view (i.e., by consuming our attention emotions can promote the search for reasons that justify our evaluative judgments, thereby promoting understanding of our evaluative judgements). But this is only one way in which emotions can lead to the acquisition of understanding. In this and the next subsections, I will describe another way in which emotional episodes can involve the acquisition of understanding. In contrast with Brady's and Elgin's view, my theory does not rely as heavily on the cognition component of emotions (i.e., on the fact that emotions can be sources of information of the environment), but rather on the *personal evaluation, motivational,* and *feeling* component. Ultimately, my main goal is to apply this view to two ways in which moral understanding can be acquired via emotional episodes, which I will develop in chapters V and VI.

Here are some ways in which, given their components of *cognition, evaluation, motivation* and *feeling,* emotions can lead to the acquisition of understanding, in light of the *curious botanists'* example. *Curiosity* will be the exemplary epistemic emotion since it clearly illustrates the way in which an emotion directs attention. I will set aside the components of intentionality and dynamism, given that they do not seem to play a substantive role in promoting understanding.

The *cognition* component of emotion can supply 'the required information about a given situation' (Ben-Ze'ev 2010: 47). As Elgin puts it, if our emotional responses accurately correlate with the events or objects that trigger them, we can use them as *sources of information* about the environment (2008: 35). The botanists' response of curiosity towards the new species of orchid supplies information about it (e.g., its colour, its resemblance to other orchids, where does it grow, etc.). This information constitutes a certain amount of knowledge about orchids, which as seen above constitutes an instance of understanding about orchids. Although as described above, this information might be distorted given the related features typical of emotions (i.e., partiality, closeness, and an intense feeling dimension), as Ben-Ze'ev notes, 'No emotional attitude toward something can emerge without some information about it.' (2010: 47)

Even if this component of emotion is extremely helpful in the acquisition of understanding, it only provides the minimal level of understanding, that is, understanding of the features that the objects of our attention seem to possess. It is the level of understanding that Brady's theory seems to provide, in terms of assessing the accuracy of the evaluative judgements associated with our emotional responses. Ultimately, Brady's theory provides understanding of the reasons that justify our evaluative judgements, which reduces to the features of the situation or object that triggered our emotional response, and that we identify as 'genuine reasons'. In contrast, the focus of my view is not finding reasons that justify our evaluative judgements. My view aims to account for the ways in which emotional episodes can increase our moral understanding-why. I will dedicate a subsection for the rest of the components of emotion,

given that each one of them can perform an epistemic function that can lead us to acquire at least two forms of moral understanding. In chapter V I will explain in which way emotional episodes can lead us to acquire the highest level of moral understanding, and in chapter VI I will elucidate in which way emotional episodes can lead to a change in moral perspectives, thereby improving our moral understanding.

iv.ii Personal Emotional Evaluation, Attention, and Understanding

Recall that my view of emotions fits into the category of *hybrid evaluative-feeling* theories of emotion. As discussed in Chapter III, I am committed to the claim that every emotion entails a certain evaluation. It is easy to see that an emotional reaction is a mental state that involves an evaluation, given that we are not usually 'moved' towards something we are indifferent to (Ben-Ze'ev 2010: 48). Emotional reactions have a significant *personal stake* (Ibid).

The evaluative component of emotions is the component that best explains the fact that emotions direct and fix our attention. Although emotions typically have certain patterns of salience (e.g., guilt makes salient an offense we committed towards others, whereas a certain form of anger makes salient an offense that others have committed against us), fear makes salient a potential threat, etc.), what we emotionally react to is influenced by what *interests us* (i.e., what we care about).

In other words, given that the botanists are *interested* in orchids (or plants and flowers in general), they *feel curious* about the new species of orchid, and consequently, they attend to it. William James has also described the link between attention and interest:

> Millions of items of the outward order are present to my senses which never properly enter into my experience. Why? Because they have no *interest* for me. *My experience is what I agree to attend to.* Only those items which I *notice* shape my mind—without selective interest, experience is an utter chaos. Interest alone gives accent and emphasis, light and shade, background and foreground—intelligible perspective, in a word. It varies in every creature, but without it the consciousness of every creature would be a grey chaotic indiscriminateness, impossible for us even to conceive. (1890: 402-403)

If the things that we attend to are the things that interest us, when emotions direct our attention, they are directing it towards that which interests us. In other words, what interests us is what is usually described as *relevant* or *worthy of notice.* This point is key for the way in which emotions can increase our level of understanding. Not only does our emotional reaction involve either a *positive* or *negative* evaluation (Ben-Ze'ev 2010: 48), which allows to understand the nature of the object of our evaluation

better, but it can also allow us *to understand why the object of our evaluation is worthy of notice*. This can be the case even if the evaluation is negative (i.e., if it is about something that interests us even if it is regarded as something negative such as acts of discrimination).

For example, when the botanists' curiosity directs their attention to the new orchid all of the emotion's components are at play: the cognitional aspect informs them of the general features of the orchid such as colour and size, they evaluate the discovery of the orchid as positive, they are motivated to keep on studying this new kind of orchid, and suppose that they experience a pleasant feeling. However, the valence of the evaluation (positive in this case) reinforces their interest in the orchid and its features, which in turn increases their degree of attention. As James pointed out, the greater the interest the greater the attention paid, and the greater the attention paid, the greater the 'mastery' or understanding (1890: 424). Hence, this reinforcement of interest provided by the evaluative component of emotion (in this case of curiosity), at the same time involves a further level of the botanists' understanding of the orchid (i.e., to what species of orchids is more similar to, in which countries it can be found, etc.).

The reinforcement of interest provided by the evaluative component of emotions therefore increases understanding of the object of evaluation. The botanists' case of curiosity illustrates a case of voluntary attention. However, in the following chapters I will focus on involuntary emotional episodes that involve involuntary attention, and explain the creation of a new interest (and which in turn involves a higher level of acquisition of understanding). This will be the case I will mainly be concerned with when I describe the acquisition of two forms of moral understanding, even if the voluntary case can also involve the acquisition of these two forms of moral understanding. I will explain these distinctions further in chapter V.

iv.ii Feeling, Attention, and Understanding

In this subsection, my aim is to highlight the fact that the feeling component of emotions plays a significant role in the focus (or 'consumption' to use Brady's terminology) of attention, and hence in facilitating understanding.

As described in chapter III, the feeling component refers to the distinctive phenomenology involved in emotional episodes (Ben-Ze'ev 2010: 49). There is something it is like to experience different emotions (e.g., experiencing curiosity is usually pleasant whilst experiencing anger is usually unpleasant). As Ben-Ze'ev notes, it is not easy to identify the varying characteristics of the feeling component:

> 'No doubt feelings have intensity, duration, and some have location as well; but what about other qualities? The qualities of being painful or pleasurable are obvious. Some level of

> pleasantness or unpleasantness, albeit often of low intensity, is experienced by most people most of the time. In addition to pleasure and displeasure, the continuum of arousal may be a common aspect of the feeling dimension.' (Ibid: 49)

Whether the feeling that comes with an emotional episode is pleasant or unpleasant is associated directly with the evaluation component; this is a reason why it is possible to make distinctions among emotions (i.e., positive emotions are usually associated with a pleasant feeling whereas negative emotions are usually associated with an unpleasant feeling) (Helm 2009: 250).[83] Regardless of the pleasantness or unpleasantness related to the evaluation, the intensity of the feeling will have an effect on arousal (and/or cortical arousal, as explained by LeDoux 1996). In general, the higher the intensity of the feeling, the higher the arousal, and hence the higher degree of attention that the emotion produces (LeDoux 1996: 290).[84] Of course, as discussed above, the higher the degree of attention, the higher level of understanding that can be acquired.

iv.iii Emotional Motivation, Attention and Understanding

My aim in this subsection is to elucidate another kind of understanding provided by the motivational component of emotion. Given that they have different action tendencies (Gibbard 1990), emotion can direct our attention to different possible courses of action. In a case like the botanists', their *curiosity* can lead them to investigate about new types of orchids. In other words, their curiosity might direct their attention to possible actions will most likely increase their understanding about various types of orchids. The botanists' investigation can consist in going to different places, read different books, and talking to other botanists. In other words, the motivation caused by curiosity can lead the botanists to increase their understanding. Call this kind of case *emotional motivation*.

Roughly put, the motivational component of emotions involves a desire or readiness to act (Ben-Ze'ev 2010: 49). Hence, this component refers to the practical aspect of emotions. According to Ben-Ze'ev, emotional motivation is a readiness to maintain or change present, past, or future circumstances (Ibid: 49). For example, in some emotions such as *anger* the desire to act is typically manifested in presently observable behaviour. In other emotions such as *hope*, the desire to act is less evident and it usually aims at performing actions in the future. Also, according to Ben-Ze'ev, this readiness to act is directly

[83] For Helm, *danger* is what pains us when we feel fear, and *success* (say) would be what presents itself as pleasurable when we feel joy, for example (2009: 250).
[84] However, there might be cases in which increased arousal does not seem to produce a strong focus of attention. For example, in a case of 'blind panic', the intensity of the panic will likely increase cortical arousal but might not direct the attention's focus on anything in particular.

tied to the positive or negative evaluation that is also part of the components of emotion (Ibid: 49).[85] Hence, a positive or negative evaluation towards a situation can also entail taking action or being disposed to act in a way which is compatible with the evaluation. For example, Mary's anger leads her to evaluate Bob's behaviour as negative, and so it might also lead her to perform an action typically regarded as negative (such as disproportionately insulting Bob for forgetting her birthday). Since anger would not be categorised as a typically epistemic emotion, then it is hard to see how such a case of anger would lead to understanding.

Like in the case of epistemic emotions (see chapter III), I argue that typically moral emotions can motivate understanding. In the case of an emotion such as *guilt*, for example, the desire to make amends can involve wanting to change the past, the present, or the future. However, in this case, the evaluation of the emotion is not compatible with the action that the emotion tends to motivate. For example, Mary's guilt for having insulted Bob leads her to evaluate her behaviour as negative. Nonetheless, the action that she decides to take in order to make amends for her offence is typically regarded as positive (such as apologising to Bob). Consequently, the motivation to apologise to Bob caused by guilt may lead Mary to increase her understanding why insulting people is *prima facie* morally wrong (e.g., because it makes her feel like an unfair friend). Additionally, Mary might increase her understanding why apologising to Bob seems *prima facie* morally correct (e.g., because it makes both her and Bob happy).

Conclusion

In this chapter, I have argued that given attention's features of selectivity and clarity (James 1890) exercising attention is a necessary condition to acquire understanding-why in general. I have also argued that emotions can focus our attention in a special way, and hence aid our acquisition of understanding-why. Although desires could also be regarded as mental states that direct our attention and allow us to acquire understanding in this way, I argued that the necessary feeling component of emotions, and the fact that they always involve a lack of rational control makes emotions more useful for the acquisition of understanding.

In subsection i.ii, I mentioned that the involuntariness of attention is particularly useful in the sense that it can lead to the acquisition of understanding that we would otherwise not get, and this kind of attention can be provided by emotional episodes. The importance of this feature of emotions will be elucidated in the next two chapters (particularly in chapter VI). I suggested that, given their components and the

[85] Bennet Helm illustrates nicely how the motivational component of emotion relates to attention. For Helm, emotions are 'intentional feelings of import' (2009: 250). He says: 'For something to have import to you—for you to care about it—is (roughly) for it to be worthy of attention and *action*.' (Helm 2009: 252, emphasis mine).

way in which they can direct our attention, emotions can also allow us to acquire moral understanding. I will develop this claim in the next two chapters.

Chapter V

Emotional Acquaintance

In chapter I, I suggested that—given her emotionless moral training in the moral laboratory—Mary was lacking at least one form of moral understanding: the highest level. As I argue in this thesis, the highest level of moral understanding consists in part in an adequate *first-hand emotional experience* with tokens of morally appraised types of actions, and which I have yet to describe. In chapter II, I argued that the relevant problem with deferring to the moral testimony (i.e., pure and impure testimony about what is putatively morally right or wrong) of others was epistemic, given the absence of the aforementioned first-hand emotional experience. In chapter III, I described the components of emotion (i.e., intentionality, cognition through patterns of salience, evaluation, motivation, and feeling), and pointed out that each emotion tends to make certain things salient (e.g., *shame* makes salient some personal inadequacy or inadequacies). Lastly, in chapter IV I argued that by focusing our attention, emotions can help us to acquire understanding of the objects that we evaluate (e.g., the botanists' curiosity motivates them to study the new species of orchid).

In a nutshell, I argued that emotions can allow us to acquire understanding in the following four ways: (i) by making salient different aspects of the objects that we evaluate and thereby allowing us to acquire morally relevant information about these objects (i.e., information relevant to the correct moral appraisal of these objects), (ii) by distinctively focusing our attention on the objects that we evaluate and thereby allowing us to assess whether we regard these objects as positive or negative, (iii) by motivating us to perform actions that will either increase our ability to respond appropriately to the evaluation of an action, or by prompting us to keep on understanding the objects of evaluation, and (iv) by allowing us—through feeling— to understand what it is like to experience different objects of evaluation.

As mentioned in the previous chapter, these claims do not yet show that emotions are necessary for the acquisition of any form of understanding. Emotional episodes can sometimes mislead us and impair our understanding, as it was emphasised in chapter III. Rather, as argued so far, my view is that at least in some cases emotions can be *epistemically beneficial*. In other words, in some cases emotions can aid understanding in general. However, in this fifth chapter I will argue that the highest level of moral understanding *necessarily* requires an *emotional acquaintance* with morally appraised actions.

In the first section, I will begin by elucidating the view that moral understanding comes in degrees (Hills 2009: 103; Sliwa 2017: 537-548). I will do this by referring back to the case of Moral Mary, and by describing what would constitute the different levels of the different putative instances of moral understanding (e.g., deferring to the pure moral testimony of others would be the lowest, and *virtuous emotional acquaintance* would be the highest). Secondly, I will point out that it is possible to acquire

understanding of morally appraised actions in multiple compatible ways. For example, via testimony, moral reasoning, emotional experiences, performing *prima facie* moral actions, via imaginings, epiphanies, contemplation of a work of art, or through the creation of artistic works, scientific discovery, etc. Thirdly, I will emphasise that a thorough account of moral understanding should consider upstream (i.e., non-practical), downstream (i.e., practical) and combined (i.e., both practical and non-practical) instances of moral understanding, and argue that emotional episodes are epistemic mechanisms (i.e., mental modes related to the acquisition of understanding) that can provide us with these different instances of moral understanding.

In the second section, I will describe three ways in which emotional episodes can provide us with upstream and downstream (and both) instances of moral understanding using Mary's experience once she leaves the moral laboratory as an example. Again, I plan to show that given their components and the way in which they can focus our attention, emotions (in this case typically *moral emotions*) can be very useful for acquiring moral understanding. However, it seems that Mary can acquire the same instances— and probably many more— of moral understanding in the absence of emotion once she has left the laboratory. Still, I will argue that a first-hand emotional experience (i.e., emotional acquaintance) with morally appraised actions can be a *distinctive* way of acquiring moral understanding, as opposed to non-emotional first-hand experiences.

In the third section, I will describe what virtuous emotional acquaintance consists in. I will understand virtue as the ability to '…recognize requirements which situations impose on one's behaviour' (McDowell 1979: 333), and rely on the Aristotelian model of the emotionally virtuous agent (*NE*, II, V: 1106b). Given that the emotionally virtuous agent experiences emotions '…at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way,' virtue will be necessary for acquiring correct moral understanding of putative moral actions (i.e., virtue will be necessary to avoid misfiring emotional episodes) (Ibid: *NE*, II, V: 1106b). Besides the valuable upstream and downstream moral understanding that emotions can allow us to gain due to their components, virtuous emotion is also epistemically privileged given the *perceptual awareness* that comes with it (Starkey 2008: 425). I will then describe the perceptual awareness of the morally relevant features of a given situation involved in virtuous emotion, and argue that putatively moral actions are understood to their highest degree by emotional *acquaintance* (e.g., Mary fully understands the wrongness of lying once she experiences different episodes of *guilt* related to lying). I will do this by defining acquaintance as a "way in which the mind can supposedly be 'directed' at an object, as a genuine relation between something actual and a subject" (Raleigh 2019:2), and by comparing acquaintance cases of emotionless and emotional Mary, virtuous but emotionally deficient Mary, and her emotionally virtuous sister Jane.

Ultimately, the upshot of this thesis is to show that, all things being equal and given the components of emotions as well as the fact that they can optimally focus our attention on the objects that we evaluate, emotional acquisition of moral understanding is *qualitatively* better epistemically than the acquisition of moral understanding without emotion. In particular, the *phenomenological experience* provided by the *feeling* component is necessary for acquiring full understanding of evaluative actions such as actions that we regard as moral, for the reason that this is the way in which— as human beings—we are capable of experiencing such actions. Cognition, evaluation, and motivation can be provided by other mental states—such as desires—but only emotion provides them in *conjunction* with feeling. Before experiencing *guilt*, for example, Mary did not know *what it is like* to wrong others—at least in all of the dimensions of feeling, i.e., physical and psychological and this sole fact already made both her capacity for moral understanding and her already-acquired instances of moral understanding *deficient*. I suspect that this claim regarding the privileged phenomenology that emotional episodes provide applies to the acquisition of most kinds of understanding, but here I am only committed to defend the claim for moral understanding.

i. The Degrees of Moral Understanding

In chapter I, I told the story of Moral Mary, who spent the first eighteen years of her life at a moral laboratory where she was emotionally 'sedated', and where she was trained to acquire as many instances of moral understanding of actions labelled as 'right' and 'wrong' as possible. Mary was also expected to make use of her moral understanding once she left the laboratory, and once she recovered her capacity of experiencing emotions.

Mary's moral training can be summarised thus, she:

Read about the history of Western morality.

Deferred to the pure and impure moral testimony of her parents.

Deferred to the pure and impure moral testimony of her teachers and the supposed moral experts.

Reflected about ethical and metaethical theories, and ethical dilemmas.

Was shown films and images about putative morally permissible and impermissible actions.

Learned to formulate positive and negative judgements about the actions that were shown to her.

Learned to recognise and distinguish putatively morally right actions from putatively morally wrong actions.

Was asked to provide autonomous explanations of her moral verdicts.

Was capable of drawing distinctions among certain right and certain wrong actions.

Was encouraged to imagine possible answers to different moral dilemmas.

Depending on how one decides to define moral understanding—for example, as a list of six reasoning abilities that account for instances of understanding-why (Hills 2009), or as a compilation of various ways in which an agent is capable of acquiring moral knowledge (Sliwa 2017)—the assessment of whether Mary came to *understand all there is to understand* about the putative moral statuses of various tokens and types of action can vary. Supposing that inside the laboratory, Mary did possess the six reasoning abilities[86] that Hills has argued are required for moral understanding to the greatest extent (just like the supposed moral experts who taught her did), it is possible to conclude that she indeed acquire the highest level of moral understanding. Hills says:

> 'You have minimal moral understanding if you correctly believe that q is why p and you can follow an explanation of why p. You have greater understanding the more you fulfil i–vi, and you have full understanding if you have i–vi to the greatest extent.' (2009: 103)

On the other hand, according to Sliwa's Moral Knowledge Account (MKA), the degree to which Mary would understand why certain actions are right or wrong would correspond to *how much* she knows about why these actions are either right or wrong (2017: 537). Sliwa notes:

> 'Plausibly there are many distinct faculties and cognitive mechanisms by which we acquire moral knowledge: perception, imagination, intuition, our affective responses, and moral reasoning can all be sources of moral knowledge. On the Moral Knowledge Account then, moral understanding is realized by a set of different faculties and cognitive abilities. Agents can have these cognitive abilities to different degrees. And so, what exactly grounds the capacity of moral understanding may vary from agent to agent.' (Ibid: 548)

On Sliwa's reductionist account, if Mary possesses cognitive abilities that give her a lot of knowledge why certain actions are either right or wrong—and therefore understands why they are right or wrong to a great extent—then Mary's degree of moral understanding inside the moral laboratory must be very high, if not the highest.

Hence, in light of both accounts it is possible to describe Mary's degree of moral understanding as *very high* before leaving the moral laboratory and before experiencing any emotions. Let us assume as well

[86] Recall that the set of abilities (where *q* is why *p*) are the following: (i) follow an explanation of why *p* given by someone else; (ii) explain *p* in your own words; (iii) draw the conclusion that *p* (or that probably *p*) from the information that *q*; (iv) draw the conclusion that that *p'* (or that probably *p'*) from the information that *q* (where *p'* and *q'* are similar to but not identical to *p* and *q*); (v) given the information that *p*, give the right explanation, *q*; (vi) given the information that *p*', give the right explanation, *q*'. (Hills 2009: 102-103)

that Mary's capacity for moral reasoning and for acquiring moral understanding is optimal. By the description of both accounts of moral understanding, it is also possible to conclude that understanding the testimony that she has received from her parents and teachers involves the *lowest level* of moral understanding; this is the case perhaps because it describes the start or the origins of moral understanding.

The final assessment of Mary's degree of moral understanding would occur once she has left the laboratory for good, and begins to experience emotions. Recall that the goals of the experiment were (i) that Mary acquired as much moral understanding as possible, (ii) that Mary was capable of making use of her moral understanding outside the laboratory, and (iii) to show the relevance of emotional episodes in the acquisition of moral understanding.

According to Hills's view, it seems that Mary can gain more moral understanding by improving her reasoning even more through the experience of new scenarios and social interactions. However, it is not clear that Hills's moral reasoning view can account for Mary's *increase* in moral understanding once she leaves the laboratory and undergoes *emotional* experiences associated with different types of putative moral actions (henceforth just 'moral actions'). Hills would not argue that before leaving the moral laboratory Mary is necessarily a good person, or that she possesses all of the virtues given her high level and capacity for moral understanding. She explains that being a good person and reliably acting rightly involves in part having our whole self (i.e., our thoughts, decisions, feelings, emotions, and actions) structured by our sensitivity to morality (i.e., the capacity to respond to moral considerations in all aspects of one's character, whenever they are relevant) (2009: 112). Given her incapacity to experience emotions, Mary could not have formed a fully virtuous character inside the laboratory according to Hills's view. According to Hills, Mary might form a virtuous character outside the laboratory, given that she may be able to orient her emotional responses appropriately. Still, it remains unclear on Hills's account how Mary's newly experienced emotions once she has left the laboratory contribute to an increase in moral understanding (i.e., it does not provide an explanation of the function of emotions, if any, in abilities i-vi).

On Hills's account, someone who is sensitive to the features of actions that determine whether those actions are right or wrong, and to whether she has reason to perform those actions is appropriately *oriented* (Ibid: 112). According to Hills, it is possible to be sensitive to those relevant features of moral actions without conceiving them in explicit moral terms (e.g., we may feel anger towards the practice of bullying without explicitly labelling bullying as 'wrong'). In fact, Hills seems to consider emotional responses as a way of recognising moral reasons (2009: 109). However, when it comes to the formation of moral judgements, *appropriate orientation* involves making use of our moral understanding (i.e., our moral reasoning abilities) (Ibid: 112). Moreover, as mentioned earlier in chapter I, Hills argues that

these abilities are necessary, and perhaps jointly sufficient for moral understanding why *p,* provided that it is true that *p,* and that *q* is why *p* (Ibid: 103).

There are two consequences—relevant to the Moral Mary case—that follow from Hills's view. The first one is that inside and outside the laboratory, Mary can possess the abilities of moral understanding to the greatest extent without *any* emotion. Although Hills would argue that possessing these abilities does *not* guarantee that Mary is a morally good and virtuous person[87] (inside and outside the laboratory), her view neglects the epistemic role that emotions can play in acquiring moral understanding. The second one is that her moral reasoning view does not account for an *increase* in the abilities or understanding that Mary could gain from emotional experiences.

On the other hand, it is also possible to deploy Sliwa's Moral Knowledge Account to try to explain the increase of Mary's moral understanding outside the laboratory. Recall that Sliwa has defined moral understanding as the ability to know right from wrong (2017: 523). Sliwa's MKA explains that moral understanding can be acquired via multiple ways (e.g., perception, imagination, intuition, emotional responses, and moral reasoning). Sliwa's MKA does take into account the possible epistemic role that emotions can play in the acquisition of moral understanding. According to Sliwa's view, Mary's degree of moral understanding would be determined by three considerations: (i) Mary's capacity of moral understanding (2017: 548), (ii) Mary's quantity of moral knowledge (Ibid: 537), and (iii) the range of circumstances for which Mary is in a position to know what the right thing to do is (Ibid: 550).

As mentioned above, Mary's capacity for moral understanding as well as her amount of knowledge acquired in the lab is very high. And if it is the case that inside and outside the laboratory, she knows what is the right thing to do in a wide range of cases, then according to Sliwa's account Mary's level of moral understanding is definitely high or very high. However, in section i.i I will argue that although Sliwa's view includes emotional responses as a way of acquiring moral understanding, it is not clear that on her view emotions are particularly important cognitive mechanisms for the acquisition of moral understanding. I will argue that they are.

Before I describe (in the next section) in what way my 'moral epistemic sentimentalism' would describe Mary's degree of moral understanding inside and outside the laboratory, it is important to point out the *phenomenological aspect* associated with the acquisition of moral understanding. In chapter IV, I mentioned that the feeling component of emotions as well as the attentional focus that emotions can provide, were the features of emotion which made emotional episodes phenomenologically distinctive. However, emotional episodes are not the only way in which acquiring understanding can involve a phenomenal aspect. Sliwa says:

[87] Hills argues that virtue 'requires that both your motivation and your judgement are responsive to moral reasons' (2009: 112).

> Many writers on understanding have noted that there's a phenomenal aspect to achieving an instance of understanding (e.g., Zagzebski 2001). This phenomenology is varied. This is reflected in how we describe our moments of moral insight: we are hit by the realization that we must help, it dawns on us (sometimes painfully) that our remark was inappropriate, we feel that someone else's action was morally wrong, we see that we must tell the truth. The Moral Knowledge Account explains why: imagination, moral perception, emotional responses, and reflection are all ways of achieving moral understanding and all have their own distinctive phenomenologies. (2017: 550)

I highlight the phenomenological aspect that can be present in the acquisition of moral understanding, given that it will be the *criterion* for placing each kind of moral understanding in a scale from lowest to highest. I will explain the reasons for this in section i.i. Below is a model of moral understanding that illustrates kinds of low, middle, and high levels of moral understanding, which I will apply to the case of Moral Mary. I will describe what the *highest* level of moral understanding would look like in section iii of this chapter.

The following suggested levels of moral understanding are not fine-grained (although I have argued that the lowest level of moral understanding is deferring to pure moral testimony, and I will argue that the highest level would be a virtuous emotional acquaintance with different moral scenarios). Nonetheless, they are based on what usually seems to happen after we are first told by our parents or the people involved in our upbringing that certain actions are morally right and certain actions are morally wrong. It seems intuitive to think that as time passes, our level of moral understanding is likely to increase as we undergo more and new experiences that involve moral evaluations; however, like in Sliwa's and Hills's account, the capacity that each agent has for acquiring moral understanding will also be a factor in determining their level of moral understanding. In other words, two agents of the same age who happen to have had very similar life experiences will not necessarily have the same degree of moral understanding, since one of them might have a greater capacity of acquiring moral understanding than the other.

i.i The Levels of a Moral Epistemic Sentimentalist Account of Moral Understanding

The Moral Epistemic Sentimentalist Account (MES) is an account of moral understanding that emphasises the importance of the role that emotions can play in the acquisition of moral understanding. As discussed so far in the previous chapters of this thesis, the claims that ground MES are the following: (a) moral understanding is reducible to moral knowledge (b) moral understanding comes in degrees (c) emotions are sources of salience and can direct our attention to the morally relevant features of morally

appraised actions, and (d) emotions can provide moral understanding in a distinctive way given their components.

MES also has a putative model of levels of moral understanding. The following description is only meant to serve as guidance to assess how low or how high is someone's level of moral understanding. The levels focus on seven types of moral understanding: (i) pure and impure moral deference (ii) moral reasoning (iii) moral judgement formation (iv) recognition of morally relevant features of the 'objects' that we evaluate (e.g., actions, situations, people) (v) experience of positive and negative emotions usually associated with moral judgements (i.e., moral emotions) (vi) moral deliberation, and (vii) moral motivation. The account is defined as 'epistemic sentimentalist' given that it seeks to explain in which ways emotional episodes can increase our acquisition of moral understanding given the components of emotion. The account seeks to show how can emotions play a positive epistemic role (i.e., further our upstream and downstream moral understanding), by being part of or exerting some influence in some of the other instances of moral understanding.

Levels of Moral Understanding

Low: the agent tends to defer to pure and impure moral testimony without much reasoning. She usually becomes aware that certain actions that she and others perform are judged as morally right or as morally wrong. She is not usually able to recognise the morally relevant features of whatever she evaluates as morally right or wrong. She typically experiences positive emotions when others and herself perform 'right' moral actions, and negative emotions when others and herself perform 'wrong' moral actions. In other words, she tends to experience so-called *moral emotions* such as *guilt, shame* (Gibbard 1990), *compassion, anger, and disgust* (Haidt 2003; Prinz and Nichols 2010). It is usually very hard for her to decide what is the right course of action when she faces difficult moral situations. She is typically unreliably motivated to do the right thing.

Middle: the agent tends to defer less to pure and impure moral testimony, and to increase her reasoning abilities. For example, following Hills' account of abilities she might possess abilities i-iii (i.e., follow an explanation why *p* given by someone else, explain why *p* in her own words, and draw the conclusion that *p* —or that probably *p*— from the information that *q*). She usually begins to judge and recognise—autonomously—certain actions as right or wrong. She tends to be even more self-aware (i.e., conscious) of the positive or negative evaluation provided by her emotional responses. She usually experiences less difficulty in deciding what the right thing to do is when she faces difficult moral situations. Her reliability to be motivated to do the right thing typically increases.

High: the agent tends to defer only to impure testimony from agents whom she regards as experts, and her reasoning abilities are very competent. She usually judges and recognises certain actions as right or wrong autonomously. She is typically constantly aware of the positive or negative evaluation provided by her emotional responses. She finds little difficulty in deciding what the right thing to do is when she faces difficult moral situations. She tends to be reliably motivated to do the right thing.

I will now assess Mary's level of moral understanding *inside and outside* the moral laboratory, in accordance with the seven types of moral understanding that I have identified at the start of this section. It will be important to keep in mind as I mentioned earlier in this section, that Mary's capacity of moral understanding is very high. Also, according to Hills's view, given that she possesses the six abilities for moral reasoning to a great extent, it can be argued that inside and outside the laboratory her level of moral understanding will most likely continue to be very high. Given that Mary also meets Sliwa's requirements for possessing a high degree of moral understanding, it can also be argued that according to Sliwa's view, Mary's level of moral understanding inside and outside the laboratory can remain high or very high. It is not clear however, how both views account for Mary's acquisition of moral understanding improving once she leaves the moral laboratory and is no longer emotionally sedated. In other words, neither view describes the distinctive role that emotional episodes play in the acquisition of moral understanding.

Inside the Moral Laboratory – According to the Moral Epistemic Sentimentalism Account

It has become clear by now that Mary's level of moral understanding inside and outside the laboratory is very high. However, *inside* the laboratory she has not experienced moral emotions—nor any emotion for that matter—which according to the Moral Epistemic Sentimentalism account is a typical aspect of low levels of moral understanding. Hence, despite her arguably high moral understanding, Mary has not experienced positive emotions when others and herself perform 'right' moral actions, nor negative emotions when others and herself perform 'wrong' moral actions. She has not become increasingly (nor constantly) self-aware of the positive or negative evaluation that tends to come with emotional responses, which is characteristic of middle and high levels of moral understanding.

Otherwise, Mary defers only to impure testimony from agents whom she regards as experts, and her reasoning abilities are very competent. She also judges and recognises certain actions as right or wrong autonomously, and finds little difficulty in deciding what the right thing to do is when she is presented

with dilemmas or difficult moral scenarios. However, it is not clear how Mary can be reliably motivated to do the right thing, since she is cut out from possible moral interactions with the outside world. Nonetheless, for the sake of the argument we can assume that Mary's formidable reasoning abilities and great quantity of moral knowledge can motivate her to reliably do the right thing inside the laboratory.

So, for example, Mary is shown a video of a bullfight. She is told by some of the supposed moral experts that bullfighting is *prima facie* morally wrong due to the fact that it reinforces cruel practices towards animals, and she defers to them. After watching the video, she can also provide other explanations that show why bullfighting is putatively wrong: it inflicts suffering on bulls (Mary understands that suffering and cruelty should be avoided), it promotes a form of entertainment that involves death, it can lead young people to believe that killing bulls in such a way is an art form thereby distorting their aesthetic appreciation, etc. Hence, she autonomously judges and recognises bullfighting as putatively wrong. She is told to imagine the complicated scenario in which her father is a bullfighter, and that therefore it is expected that she supports his profession and attends his bullfights. Mary is able to argue that in this situation the right thing to do is to refrain from attending her father's bullfight, despite it being so important to him. She also says that in a case like that she would be motivated to not attend her father's bullfight, by the reasons she has given to believe that bullfighting is morally wrong.

The previous example shows instances of Mary's understanding of the putative wrongness of bullfighting, without any emotion. I will now use the same example to show Mary's understanding of the wrongness of bullfighting, outside the laboratory when she is able to experience emotions. My aim will be to show that once Mary is able to undergo emotional episodes, her level of moral understanding will increase.

Outside the Moral Laboratory– According to the Moral Epistemic Sentimentalism Account

Recall that the view of emotion that I have assumed, explains that the components of emotion are intentionality, cognition, evaluation, motivation, and feeling (chapter III, section I). The intentionality component only elucidates that when we experience emotions, these have an object (e.g., disgust *for* bullfighting). Shortly, I will explain how do the other components of emotion can increase Mary's level of understanding of the putative wrongness of bullfighting.

As mentioned before, Mary has already deferred to the impure testimony of the supposed moral experts, who told her that bullfighting is morally wrong given that it reinforces cruelty towards animals; she can provide explanations that show why bullfighting is wrong, and she can autonomously recognise and form the judgement that bullfighting is wrong. This time, however, she does not have to watch a video

of a bullfight nor she has to imagine the scenario where she is expected to attend her father's bullfight. Mary goes to a bullfight for the first time, and experiences *disgust* and *compassion*. In which ways does experiencing such emotional episodes increase her moral understanding?

Although Mary had already watched the bull's suffering during a bullfight, her compassion and disgust provide her with more relevant information about the bull's suffering: she fixes her attention on the way the bull moves uncomfortably, on the loud noises that it makes, on how the blood keeps flowing from its back. Before experiencing these emotions, she had not noticed another relevant element which promotes animal cruelty: the blind-folded horses that are used in the bullfight might get killed or hurt by the bull. The negative evaluation that her disgust and compassion provide her with is phenomenologically distinctive: she has further thoughts about animal suffering (and other forms of suffering), which also makes her feel sadness; she also feels 'a knot' in her stomach. Her previous conviction on her judgement on the wrongness of bullfighting is reinforced, and together with her previous motivating reasons her disgust motivates her more radically to never attend a bullfight again. Moreover, her disgust and compassion now motivate her to protest again bullfighting.

Mary's previous emotional experience was—at least in great part— *epistemic* given that it increased her understanding of the wrongness of bullfighting, and this in turn led to a general increase her general level of moral understanding. Mary acquired *new relevant information* of the act of bullfighting, her negative evaluation of bullfighting was *reinforced*, she experienced *a certain feeling* towards bullfighting which also led her to form new thoughts, and she was motivated to act in a way *she had not acted before*. Roughly put, she increased her amount of knowledge, she is now able to provide new reasons and explanations why such practice was wrong, she now associates a certain feeling to bullfighting (which also led her to new ways of thinking about the wrongness of such actions), and she is now motivated to act in certain ways. According to the Moral Epistemic Sentimentalism (MES) that I will defend, this is what an emotional acquaintance with a type of moral action would look like, although I will provide a general definition in section iii. In the next subsection, I will describe in more detail the phenomenology of Mary's emotional experience and the relevance that it has on the acquisition of moral understanding.

i.ii The Relevance of the Phenomenology of Mary's Experience Outside the Laboratory

There is a distinctive phenomenology of emotional episodes that leads to an increase in moral understanding. Although the phenomenological experience of every moral emotional acquaintance will tend to vary from agent to agent, there are two general identifiable aspects of the phenomenology of these emotional episodes: *self-awareness* (i.e., consciousness) and *feeling*; both aspects can allow us to

distinguish our emotional experience from other mental states, such as desires. As mentioned in chapter IV, the cases of attention produced by emotions that I am interested in are those cases in which we are self-aware of the object of our attention, and of our perception of the object. When Mary experiences compassion and disgust the first time that she attends a bullfight, all things being equal, she is at least self-aware of the new understanding that she is acquiring, as well as of the feelings that she is experiencing; as Ben-Ze'ev says, there are no 'unfelt feelings' (2010: 49).

Feeling

In chapters III and IV I have described feeling as the component of emotions which provides their distinctive phenomenology (Ibid:49). Feelings usually have intensity, duration, location, and the qualities of being painful or pleasurable (Ibid:49). I have also emphasised in this and the previous chapter that feeling is essential to emotion, and that it is in part what distinguishes emotions from other mental states. The feelings that we tend to experience when we undergo emotional episodes can be bodily (e.g., to blush, to have a stomach-ache) and/or psychic (e.g., being excited, being interested) (see Stocker 1983).

Goldie argues that emotions usually involve two kinds of feeling: *bodily feelings* and *feelings towards* (2002). As explained in chapter III, in this thesis I am not endorsing Goldie's view of emotion, but his description of feelings is useful to elucidate the phenomenology of emotional episodes. According to Goldie, a bodily feeling is a sensation directed towards one's body, as being a certain way; for example, when you feel an excruciating pain in your elbow, the object of the sensation is your elbow which feels a certain way: excruciatingly painful (2002: 236). Bodily feelings, though, are not necessarily experienced when we experience emotions. On the other hand, we tend to experience feelings towards when an emotion is directed towards an object: a thing, a person, a state of affairs, or an action or event; for example, when you are upset at the way she purposedly turned her back on you when you came into the room, the object of your emotion is that action (Ibid: 241).[88] When undergoing an emotional episode, we can experience both kinds of feeling, directed both at our body and at other objects. The phenomenology of feeling will be useful to further describe Mary's moral emotional episodes after she

[88] Goldie describes 'feeling towards' as 'an unreflective emotional engagement with the world beyond the body; it is not a consciousness of oneself, either of one's bodily condition or of oneself *as* experiencing an emotion. Feeling towards is this something that a creature which is incapable of self-reflective thought – a dog or a toddler, for example – could achieve. We adult humans, however, are capable of a turn of reflectiveness: we are capable of *noticing* that we have feelings towards something.' (2002: 241-242) Although my view is also aware of the possibility of unreflective emotional engagement with objects, it focuses on the cases in which there is a self-aware emotional engagement (even if it is not overly reflective, or reflective at all for that matter), due to the fact that the act of self-awareness of moral emotional episodes in itself can lead to an increase in moral understanding. I am assuming here that not all acts of self-awareness necessarily involve conscious reflection.

has left the laboratory, and it will become clearer after discussing the following 'before' and 'after' example.

In his paper entitled "Psychic Feelings", Michael Stocker says:

> "Having fallen on the ice, the very same knowledge of (and wishes to avoid) the dangers of walking on the ice are "emotionally present" to me. They concern me to the point of my being afraid. Before the fall, I had only an intellectual appreciation of the very same dangers (and a rather pro-forma desire to avoid them). Then I only saw the dangers, now I also feel them." (1983: 20–21)

It is easy to compare the event of 'seeing dangers' before 'feeling them', to Mary's situation inside the moral laboratory. In his example, before falling on the ice, Stocker had an 'intellectual appreciation' of the *dangers* of falling on ice; however, once he experienced walking on ice and falling, his fear led him to further appreciate or acquire a *further understanding* of the dangerousness of falling on ice, captured by the phenomenology of his fear (i.e., he understands now how these dangers *feel*). Analogously, before Mary attended a bullfight —and without any emotion—she already understood the wrongness of bullfighting. However, after she witnesses a bullfight and she is able to experience compassion and disgust, she now understands the wrongness of bullfighting better, or to a greater extent (in Stocker's words, she 'feels' the wrongness). Stocker's reasoning applied to 'feeling rightness and wrongness' mirrors Hume's famous claim that morality is more 'properly felt than judged' (T 3.1.2.1).

Peter Goldie also explains the difference between before and after 'seeing' and emotionally experiencing the dangers of walking on ice. Goldie says:

> 'When we think of something as being dangerous, we might just think of it as meriting fear, and we can do that without actually feeling fear towards it. Then, when we come to think of it with fear, the dangerousness of the object, and the determinate features towards which the thought is directed, is grasped in a different way. That is to say, the content of the thought is different; one's way of thinking of it is completely new. It is not just the old way of thinking of it, plus some new element. Rather, it is more like coming to see a hidden shape in a drawing, or coming to see the shape of the face on the visible surface of the moon: one's way of seeing is completely new.' (2002: 243)

Goldie emphasises that what is gained after experiencing a feeling towards (i.e., an emotion) is a new *way of thinking* of something that we were already able to think of in another way. Goldie has his own thought experiment to describe this newly acquired way of thinking via emotion. He tells the story of

Irene, the ice-scientist. Irene has 'complete knowledge of the dangers that can arise from walking on ice' (Ibid: 244-245). Irene is also 'icy-cool', and has never experienced *fear* (she has been brought up in a very pampered way). Nonetheless, she has both a theoretical concept of dangerousness, and a theoretical concept of fear (e.g., she knows that people are usually afraid when they perceive dangerous things). One day, Irene falls on the ice and for the first time she feels fear towards it, and Goldie explains that she now knows 'from the inside' *what it is like* to feel fear; and so, Irene has gained a new concept—a phenomenal concept. Additionally, Goldie argues that she has acquired a new perceptual concept of dangerousness: when Irene now thinks of the ice as dangerous, she can do so *with fear*. Goldie says: 'Before, she knew that the ice was dangerous, for she knew that it merited fear, but, because she now is able to think in a new way of fear, *she now understands in a new way what it is for the ice to be dangerous*.' (Ibid: 245, emphasis mine).

Just like Goldie's Irene after she has fallen on ice thinks of its dangerousness in a different way, when Mary leaves the moral laboratory and feels *compassion* and *disgust* towards bullfighting for the first time, she then thinks of the wrongness of bullfighting in a new way. Thinking of bullfighting in a 'new way' can involve, as Goldie suggests, noticing new things about it; as I mentioned before, when Mary finally attended a bullfight, she fixed her attention in some aspects she had not noticed before, such as the agonising noises the bull makes, and the fact that the blindfolded horses are also in great danger. After this experience, it is possible to argue that Mary has acquired the phenomenal concepts of what it is like to feel compassion and disgust as well as the concept of the wrongness of bullfighting (i.e., she can think of bullfights being wrong with disgust and compassion).

In sum, other things being equal, it is possible to argue that when she attended a bullfight for the first time and she was no longer emotionless, Mary's phenomenological experience of disgust and compassion allowed her to increase her understanding of the wrongness of bullfighting. So far, this phenomenology can be described as experiencing bodily feelings, and feelings that can lead to the acquisition of new phenomenal and perceptual concepts. Below I continue to describe this phenomenology in terms of self-awareness.

Self-awareness

Part of the phenomenology of certain emotional episodes consists in the self-awareness of attentional focus. In chapter IV, I argued that my view concerned instances where attention involves consciousness, and took consciousness to be a *self-aware mental state* whereby agents are both aware of the object that they are perceiving, and aware of the activity of their perceiving the object (Brentano 1874). Hence, when Mary's emotional episode of compassion and disgust fixes her attention on the ways in which the

bull is suffering, she is both self-aware of the state of affairs that she is witnessing as well as of her undergoing an emotional episode. This aspect of Mary's emotional experience is also relevant, given that the before-after difference for her is not just a difference in her way of thinking of the wrongness of bullying, but also in the *effects* of these new ways of thinking. As I mentioned earlier in section i.i, Mary acquires new relevant information about what grounds the wrongness of bullfighting, she reinforces her negative evaluation of it, she is motivated to act in ways[89] she had not previously acted before,[90] and these effects in turn increase her moral understanding. Of course, it is possible to conceive of these effects taking place at an unconscious level. However, I argue that if Mary is self-aware of the increase in her moral understanding, then she will be able to make more use of it.

It is possible to elucidate the effects of Mary's bullfighting emotional episode further, by looking again at Goldie's Irene. Goldie argues that besides her new way of thinking of fear and dangerousness, Irene also gains 'new powers and potentialities of thought, imagination and feeling.' (2002: 245) Goldie explains that Irene can deploy her new phenomenal concept by remembering experiences of danger in a fearful way that she was unable to in the past, and when she imagines someone else feeling fear, she can imagine what it would be like to be in their shoes (Ibid: 245-246). Similarly, Mary can now associate her phenomenal concepts of compassion and disgust with new experiences of wrongful actions, and now when she imagines someone else feeling compassion and disgust, she can imagine what it is like for them to feel these emotions. In section iii, I will describe in more detail Mary's first-hand emotional experience at the bullfight—and the abilities to remember and to imagine— which she probably gained from it (Lewis 1988).

In sum, the phenomenology of Mary's new experience of compassion and disgust plays a crucial role in Mary's process of acquiring more moral understanding about the wrongness of bullfighting after she has left the moral laboratory. If Mary's emotional experience increased her moral knowledge, then—as argued by the MKA and MES— more knowledge would have led her to a higher level of moral understanding. Nonetheless, according to MES, there are also some ways in which emotions could lead Mary epistemically astray. As discussed in chapter III, emotions can misfire and lead to incorrect judgements and actions. In the next subsection, I will discuss possible problems with Mary's new emotional episodes. My aim is to show that according to MES emotional episodes will not always lead to the acquisition of moral understanding, and so it can avoid objections about the possibility of emotions impairing moral understanding.

[89] Such as *discussing* the morality of bullfighting with her father and everyone who supports bullfighting, *stopping to attend* bullfights, *investigating* about the history of bullfighting to understand it better, etc.
[90] As described by Stocker's feeling the dangers of the ice, Goldie says that Irene also acquired an emotional desire *to avoid* the dangerousness of the ice (2002: 245).

i.iii Possible Problems with Mary's New Emotional Episodes

The Moral Epistemic Sentimentalism account describes emotions as being capable of playing a distinctive role in the acquisition of moral understanding. Nonetheless, it is a view that takes into account at least some of the ways in which emotions can hinder moral understanding.

As discussed in chapter III, emotions can be variable, volatile (Elgin 2008), lead to irrational action (Hursthouse 1991), they can be regarded as a source of weakness of the will (Davidson 1970), and lead to impulsivity (Frijda 2010). Moreover, as Goldie noted, some emotions can be systematically misleading and thus very difficult (if not impossible) to correct (i.e., recalcitrant) (2008). Hence, although Moral Mary's new emotional episodes are likely to increase her level of moral understanding, there will be at least some cases in which her emotional reactions will prevent her from a correct acquisition of moral understanding, or lead her to immoral action.

Generally, the three ways in which emotions can be epistemically unhelpful are the following: (i) they can hinder the formation of moral beliefs and evaluations, (ii) they can motivate irrational and/or immoral action, and (iii) they can be systematically misleading (and sometimes extremely difficult to correct).

For example, recall that inside the laboratory, Mary already understood and judged that bullfighting was wrong. To her dismay, when she leaves the laboratory, she learns that her father is a bullfighter. However, since she is now able to feel *love* for her dad, her love might influence her previously held judgement about bullfighting. Hence, when she attends a bullfight for the first time, her love will focus her attention on her dad's skills as a bullfighter rather than on the bull's and the horse's suffering. Also, given that Mary has acquired the capacity to feel emotions, she may not know how to regulate her emotional reactions very well, and so the love for her dad might fight off the compassion and disgust that she might also experience during the bullfight. As a result, at least three things could happen: (i) Mary might question her judgement about the wrongness of bullfighting and decide to remain agnostic (ii) Mary might start loving bullfighting and judging that it is right 'just because her dad says so' or due to the many skills of the bullfighters, (iii) Mary might still believe that bullfighting is wrong but she continues to attend her dad's bullfights anyway.

It seems that in all three cases, Mary's level of moral understanding does not increase. The least bad of the three options seems to be (i), due to the fact that although she is not acquiring moral understanding about bullfighting, her level of moral understanding is not being as negatively altered like in option (ii). Recall that given her training in the moral laboratory, Mary's general level of moral understanding was very high, which entailed that she no longer deferred to pure moral testimony; however, in case (ii) the love for her dad is making her defer to him about the morality of bullfighting, which in the scale of

MES is a feature or an indication of having a low level (if not the lowest) level of moral understanding. Lastly, in case (iii) Mary's decrease or negative change in moral understanding is shown in her actions. As she learned in the laboratory, to understand that something is wrong is in part to understand that it should be avoided. And so, when Mary still judges that bullfighting is wrong but still attends her dad's bullfights she is showing—among other things, such as weakness of the will— a deficiency in balancing moral considerations.

In this example, Mary's love for her father clearly impaired her moral understanding. Her love led her astray by hindering her previously acquired ability to recognise the morally relevant features of the practice of bullfighting, by overriding other emotions that would have focused her attention on the animals' suffering, by not allowing her to acquire more moral understanding, by making her go back to deferring to pure moral testimony and by making her act against her moral judgement.

Hence, it is clear that emotions will not always lead to the acquisition of moral understanding. Emotions can lead to erroneous moral actions and to the impairment of moral understanding just like it happened to Mary in the previous example and possibly in many other ways. However, as discussed in chapter III, human beings are capable of regulating[91] their emotions (Elgin 2008: 47). It is harder for Mary to regulate her emotions given that experiencing them is new for her, but she could still refer to strategies for calibration such as attending and reflecting on her emotional responses (i.e., noticing the situations that trigger them, the orientations they give rise to, and the opinions that they generate), as well as to the arts (e.g., listening to certain music might allow her to become more aware of her emotional responses).

There is another way in which Mary could regulate her newly experienced typically moral emotions. As mentioned in chapter III subsection ii.iii, paying attention to the emotional responses of artists could be an alternative way to calibrate emotions. Musicians, painters, dancers, etc., usually have a refined sensitivity that allows them to perform their works of art. So, for example, a way in which Mary can calibrate her emotional responses of love and admiration for her dad could be by observing a professional ballet dancer. As she observes the ballerina's graceful movements, she might realise that there are other forms of art that do not involve killing animals, and which are more admirable for it. This reflection can cause her to separate the feelings of unconditional love and admiration from the act of supporting his profession. This distinction would involve a refinement of her emotional responses. In this way, her love and admiration for her dad would no longer be leading Mary morally astray (e.g., she would not continue to attend bullfights whilst thinking that it is morally wrong to do so).

Lastly, another way in which regulation of typically moral emotions can also take place via the observation and/or imitation of the behaviour of supposed moral experts (or in general of people who

[91] I take the activities of regulation, calibration, and refinement of emotions to involve the same process which amounts to the same result: transforming emotional responses in a way that improves them.

are better than us at doing the right thing). So, for example, after she leaves the laboratory Mary will have the chance to observe her virtuous sister Jane's expertise in doing the right thing most of the time. This way, even if some of her newly experienced emotions happen to be disproportionate or inadequate, by observing Jane's behaviour she might try to imitate Jane's emotional reactions which will likely be moderate and adequate. And so, after observing that when Jane feels guilty for having betrayed Bob she apologises to him, she can observe to what extent does Jane experience guilt, and imitate her practice of apologising to someone she has wronged.

The previous examples show that although emotional responses can misfire and lead to erroneous beliefs and evaluations, as well as to immoral behaviour, it is nonetheless possible to educate them. By learning different ways to regulate our emotions we might eventually benefit from the positive epistemic role that they can play. However, in chapter III I also argued that Goldie's view on misleading emotions shows the way in which emotions can be most problematic. Due to the fact that some emotions are built into our evolutionary psychology, they can motivate us to act impulsively in a systematic way (e.g., aggression, fear of strangers, jealousy) (Goldie 2008: 155). In other words, these emotions used to play specific roles historically in previous environments, and so when they are still experienced in our present environment it is *extremely difficult* to correct them (see Gigerenzer and Selten 2001). Relatedly, Peter Railton has pointed out that cases of certain mental conditions can cause systematic cognitive difficulties associated to regions of the brain where affective reactions are produced:

> Individuals who suffer depression, mania, or schizophrenia, or who experience damage to the interface between *affective* regions and the higher cortices, show systematic problems in learning deliberation, and decision making. (Railton 2014: 844, emphasis mine)

Hence, some kinds of mental conditions can also interfere in a negative way when it comes to experiencing emotions. If Mary falls in a mental state of, say, depression, then it is possible that—systematically— she does not react in any way to her experience of guilt. Perhaps Mary's depressive state leads her to think that she is devoid of feelings just like she was inside the laboratory, or maybe the chemical imbalance in her body might stop her from feeling at least certain emotions. In this case, guilt will probably not lead Mary to reflect about the wrongness of betraying Bob, nor will it motivate her to apologise to him. Therefore, it is unlikely that an episode of guilt in a case like this will lead to any kind of moral knowledge or understanding.

An example of a specific mental condition that directly affects emotions in a negative way is psychopathy. Railton says:

> Psychopathy is thought by some to involve a profound deficit in *empathy*—and perhaps specifically emotional rather than cognitive or motor empathy (61)—resulting *in impulsivity*,

> *disinhibition of violence, instrumentalization of others, and difficulty in forming stable, reciprocal relationships*. (Ibid: 844, Railton, emphasis mine)[92]

If mental conditions like psychopathy are usually innate, and affect emotional responses in a way in which they can lead to immoral behaviour, it is hard to see how certain typically moral emotions such as empathy can help to produce moral understanding. Moreover, just like some systematically misleading emotions are extremely difficult to correct, in a case like psychopathy it seems that it will be very hard to educate or regulate emotions which are not even felt. By leaving out this group of individuals, it could be argued that MES fails to be sufficiently *democratic*. I argue that a model of moral epistemology should include as many groups of individuals as possible, and so in this sense, views about the acquisition of moral understanding should be democratic as opposed to *elitist* or *exclusive*. The reason why I argue that accounts of moral epistemology should be democratic, is that anyone who is interested in gaining or improving their moral understanding should ideally be able to find theories that inform them on how to do it. These theories should be informative about how we usually learn how to perform *prima facie* right actions, in order to be of use to those who care about acquiring moral understanding.

However, although not being democratic enough would seem like a problem that MES faces, cases like psychopathy in fact strengthen the claim that grounds MES. In other words, the claim that emotional episodes can play a very important role in doing the *prima facie* morally right thing, and hence in the acquisition of moral understanding. Psychopathic individuals are commonly thought to be immoral, mainly due to their incapacity of experiencing certain emotions (e.g., *remorse* after harming others) (see American Psychiatric Association, DSM IV-TR). Ultimately, if psychopathic individuals do not care about acquiring moral understanding, they would not enter a category of a group that that MES leaves out.

The case of psychopathy also shows that it is possible to possess reasoning abilities to a great extent, and in turn possess at least some or a high degree of moral understanding. In a way, Mary was psychopathic inside the laboratory, given that she was deprived of experiencing typically moral emotions (including empathy and remorse). I have argued that after experiencing emotions, Mary's moral understanding increased, which supports the argument that emotions are extremely important for the acquisition of moral understanding.[93]

---

[92] Railton's source is Blair, Mitchell, and Blair (2005: 698–718).

[93] Also, there is something to be said about the *democratic nature* of emotions. Despite the existence of some complications in the mental processes of some individuals which impair or eradicate their emotional reactions, emotions are *widely* experienced. On the other hand, having the capability of exercising reasoning abilities to the greatest extent might seem like a more demanding requirement for moral understanding. Arguing in more detail for the democratic nature of emotions and the impact that this has on moral understanding has fallen outside the scope of this thesis, but it strikes me as an interesting and relevant avenue for future research.

In the next section, I will identify some upstream and downstream aspects of moral understanding that can be influenced by emotional episodes. My aim is twofold: (i) to elucidate how adequate emotional episodes can influence and sometimes lead to correct moral behaviour, and (ii) to distinguish moral understanding from other types of understanding acquired by emotional episodes. In chapters III and IV I described some ways in which some emotions can motivate scientific research (e.g., *curiosity*). However, the goal of this thesis is to describe the importance of some of the ways in which emotional episodes can lead us to acquire moral understanding. Moreover, the kinds of understanding that some emotions can provide us with can have either upstream, and/or downstream aspects. Roughly put, by *upstream* aspects of moral understanding I am referring to the *non-practical* characteristics of moral understanding such as sets of implicitly held moral beliefs, attitudes, emotions, etc., upon which we can act or not act. For example, holding the belief that it is wrong to betray one's friends, disapproving betrayal in general, the experience of an episode of guilt after betraying one's friends, etc. Alternatively, by *downstream* aspects of moral understanding, simply put, I am referring to the actions that result from our implicitly or explicitly endorsed moral beliefs and attitudes (i.e., the *practical* aspects). For example, if we disapprove of betrayal, experiencing guilt may lead us to apologise to one's friends for betraying them. In other words, the act of apologising would be the downstream aspect of our moral understanding about betrayal.

ii. Upstream and Downstream Aspects of Emotional Moral Understanding

In this section, I will argue that upstream (i.e., non-practical) and downstream (i.e., practical) aspects of moral understanding can be influenced epistemically in a positive way by emotional episodes. As discussed in the previous section, emotions can sometimes mislead us and hinder our moral understanding. Still, the aim of discussing the following examples is to highlight some of the ways in which emotions can be useful for both upstream and downstream aspects of moral understanding.

Emotions and Upstream Aspects of Moral Understanding

By 'upstream aspects' of moral understanding, I am referring to the ways in which moral understanding is non-practical in the sense that it involves the set of general first-order and second-order views and attitudes towards morality that we implicitly and reflectively endorse (Sinclair 2021: 194-95). Mary's general *implicitly* held belief that 'acts of charity are *prima facie* right', and/or approval of acts of charity would be an example of a *first-order* view. On the other hand, Mary's set of general implicitly endorsed

moral views and attitudes are considered to be *appropriate* on the basis of *another* set of attitudes/beliefs that she *reflectively* endorses (Ibid: 194-95). This additional set of reflectively endorsed attitudes/beliefs would be her *second-order views* (e.g., Mary's belief that everyone should show solidarity towards those in a disadvantaged situation, and/or her approval of acts of solidarity).

Once Mary starts her life outside of the laboratory, how are her set of implicitly and reflectively endorsed moral views influenced by her newly experienced emotional episodes? Take Mary's view that acts of charity are *prima facie* right; her approval and her judgement towards these acts is now at least in part grounded or accompanied by *compassion*. Compassion now focuses her attention on the morally relevant features of particular acts of charity; sometimes she even notices elements of the same act tokens that she had not noticed before. For example, when her sister Jane buys a cup of coffee for a man who is sitting and freezing in the sidewalk, Mary's heart is 'warmed'; among other things, her compassion makes her realize that mostly everyone in the street is holding a hot drink because it is snowing; she realizes that not everyone is privileged enough to buy themselves a hot drink when the weather is very cold, she also sees how the man turned from shivering to smiling, and notices again how doing something nice for others brings happiness both to the ones who benefit from acts of charity, and to those who perform such acts.

As Goldie would put it, Mary has learned what it is like to experience compassion in cases such as 'the cup of coffee incident', and a new way of thinking about the rightness of acts of charity (i.e., with compassion at least in some cases). Notice that Mary's moral understanding is both particular and general: she acquires understanding of the way in which 'the cup of coffee incident' was right, and this particular instance of understanding increases her general moral understanding of the rightness of acts of charity.

The usual result of these episodes of 'compassion-towards acts of charity', is that Mary learns more information that grounds her belief that acts of charity are right, and the positive evaluation and feeling experienced towards this type of actions reinforces her beliefs and attitudes of their appropriateness. A case like this shows that the cognitional, evaluative and feeling components of emotion can have a positive influence in the upstream aspect of moral understanding. This positive influence can be explained in *epistemic* terms, given that her episode of compassion during the cup of coffee incident led Mary to understand the *prima facie* rightness of acts of charity better.

Emotions and the Downstream Aspect of Moral Understanding

I take the downstream aspect of moral understanding to be the *practical* aspect of moral understanding. The downstream aspect of moral understanding would be the actions and tendencies to feel (i.e.,

attitudes) that stem from the upstream aspect of it (i.e., our set of first and second-order moral views). I argue that the component of emotions that plays a crucial role in the causation of moral behaviour is the motivational component.

Hence, if Mary believes that acts of charity are right and she approves of them, it is likely that she will perform acts of charity at least in some cases. As a result of her episode of compassion during the cup of coffee incident, Mary now feels motivated by her desire to imitate Jane's generous disposition and charitable behaviour in the future. Also, thinking and remembering about the incident with compassion leads her to imagine what it must feel like to be as cold as the man shivering in the street, which motivates her to be disposed to help others in a similar situation whenever she can in the future. In other words, Mary's actual performance of different tokens of charitable actions will have been—at least partially—motivated by her episode of compassion.

Even if inside the laboratory Mary could have been *solely* motivated by certain reasons to perform charitable actions (e.g., it helps to bring about social justice, it increases happiness), or by her moral judgement itself, it is clear that her motivation to perform such actions is *improved* by her emotional episode of compassion outside the laboratory. By being motivated to imitate Jane's generous actions, she can learn how and when to perform charitable actions (and even develop a more generous disposition to act). Moreover, it can be argued that the motivational power of emotions is more direct than motivation by judgement.[94] Due to their involuntary and automatic nature (see chapter III section I, and chapter IV subsection iii.iii), emotions are more likely to prompt to action, rather than judgements. As Tappolet has explained, emotions are world-guided in the sense that they *automatically* respond to environmental states of affairs or stimuli (2016: 19-20). In other words, there is usually less deliberative or rational control when our emotional responses lead us (or make us) perform certain actions.

In sum, Mary's new capability of experiencing emotions is likely to have a positive epistemically influence in both the upstream and downstream aspects of moral understanding. Emotions can improve her capability of acquiring non-practical moral understanding, as well as her capability of acquiring moral understanding through action, and thus improve her general level of moral understanding. So, for example, when Mary experiences guilt after having insulted Bob when he did not deserve it and guilt leads her to imagine what it would be like to be insulted by Bob undeservedly, she is gaining some understanding of the wrongness of insulting someone who does not deserve it (for instance that she

[94] Prinz and Nichols explain that motivation internalists sometimes assume that moral judgements are constituted, at least in part, by emotions (e.g., the mental state of judging that killing is immoral is constituted by a mental representation of killing along with an emotional stated directed toward that represented action). Motivation externalists argue that judgements can be made without motivation but they usually agree with the internalist claim that when motivation accompanies a judgement, it is obtained from an emotional state. In other words, internalists and externalists often agree that emotions contribute to the moral motivation of those individuals who are usually motivated to act in accordance with their moral judgements. In sum, emotions can motivate us to act morally, and they can do so without a held moral judgement or as a result of a moral judgement (2010: 113-114). Regardless, it seems that in general emotions seem to provide a stronger motivational force than judgements.

would not like to be treated in such ways). Additionally, by motivating her to perform compensatory actions for her bad behaviour towards Bob, Mary also learns some or different ways to make amends, and this fact leads her to understand that trying to correct one's immoral behaviour is the right thing to do.

The role that typically moral emotions can play in the downstream aspect of moral understanding—concretely in partly producing certain kind of actions—differs from the behaviour that other kinds of emotions can produce. For example, epistemic emotions such as curiosity, usually lead to scientific inquiry, whereas guilt and other typically moral emotions like compassion concern actions that we usually judge as right or wrong, permissible or impermissible such as apologising, betraying, or helping.

Now that I have argued that emotional episodes can lead to the acquisition of moral understanding, as well as the role that they can play in both upstream and downstream aspects of moral understanding, in the next section I will describe the highest level of moral understanding.

iii. Virtuous Emotional Acquaintance

In this section, I have two aims: (i) to describe virtuous emotional acquaintance, and (ii) to compare my view of emotional acquaintance with Ballard's (2020), and explain why MES provides a more thorough account of the way in which emotions can be epistemically beneficial.

So far, I have described Mary's new emotional experiences outside the laboratory as 'emotional acquaintance' or as 'first-hand emotional experiences' with morally appraised actions. I take acquaintance to be the "way in which the mind can supposedly be 'directed' at an object, as a genuine relation between something actual and a subject" (Raleigh 2019:2). The phenomenology of Mary's new emotional experiences explains the way in which Mary can be cognitively related to moral actions after leaving the laboratory. I have argued that these new moral emotional experiences can increase her level of moral understanding. In the next chapter I will discuss the category of 'moral emotions', although, so far, I have assumed that any emotion that allows Mary to acquire moral understanding can be regarded as typically moral.

As I have argued so far in this chapter, the components of emotion can allow Mary to understand moral actions better. Before I explain why virtue facilitates and warrants a successful moral emotional acquaintance, I will emphasise the relevance of first-hand emotional experiences. I argue that it is through this kind of experience that we can understand all there is to understand about the putative

moral statuses[95] of various tokens and types of action. It is in this sense that MES explains that first-hand emotional experiences can provide the highest level of moral understanding. When Mary attended a bullfight for the first time and she experienced compassion and disgust, her understanding of the putative wrongness of bullfighting was enhanced. Roughly put, she came to understand more morally relevant features of bullfighting than she did when she could not experience any emotion. Although not all of the moral understanding that Mary will acquire outside the laboratory will be necessarily accompanied by an emotional episode, whenever she acquires moral understanding via a correct emotional episode her level of moral understanding will always improve.

In general, a first-hand experience gets us richness of content through our own perception. For example, being told that a patch of colour is red without looking at it is not the same as looking at it; in both cases we may come to know that it is red, but when we get to see it, we learn more by seeing its precise shade (Sliwa 2017:548). Similarly, as Sliwa's MKA explains, to understand why an action is wrong generally requires one to know what some of its wrong-making features are. According to Sliwa's account, Mary would also come to understand better why bullfighting is wrong once she sees a bullfight first-hand since she would access to more information that explains its wrongness; however, Sliwa's MKA does not provide a detailed explanation of the role that emotion might play in these first-hand experiences.

Sliwa does argue that certain first-hand experiences may be 'essential to gain epistemic access to certain wrong-making features because these include facts about how the action affects the wronged party: what being the victim of this wrong is like. It's plausible that you cannot fully appreciate some important aspects of the relevant experience—the shame of a survivor of sexual assault or the isolation experienced by a victim of domestic violence—unless you have undergone a similar experience.' (Ibid:549) Sliwa goes on to say that first-hand experiences can allow us to acquire particular instances of moral understanding, but that they can also broaden our capacity of moral understanding. In other words, we can gain abilities that will make us more sensitive to acquire instances of moral understanding (Ibid: 549).

Like Goldie and Sliwa, MES draws on Lewis's discussion of abilities 'to remember and to imagine' (1988) to explain the understanding that Mary acquires through her new first-hand emotional experiences outside the laboratory. Lewis's famous example to explain these abilities is the experience of tasting vegemite: 'After you taste Vegemite, and you learn what it's like, you can afterward remember the experience you had. By remembering how it once was, you can afterward imagine such an experience. Indeed, even if you eventually forget the occasion itself, you will very likely retain your ability to imagine such an experience. Further, you gain an ability to recognize the same experience if it comes again.' (1988: 17) And so, besides remembering and imagining her episode of compassion and disgust at the bullfight—thereby understanding better the wrong-making features of bullfighting—

[95] By putative 'moral statuses' I am referring to the different ways in which certain actions can be regarded as right or wrong, permissible or impermissible.

Mary can also gain the ability to recognise the ways in which in similar situations are also wrong (e.g., hunting for entertainment).

This ability and sensibility to recognise the morally relevant features of various situations is one of the main characteristics of a virtuous person. In this thesis I understand virtue in the McDowellian sense. In other words, as the ability to '…recognize requirements which situations impose on one's behaviour.' (McDowell 1979: 333) Again, inside the laboratory Mary might have had the abilities to remember, imagine, and recognise the rightness or wrongness of various situations, but without emotion her degree of virtue (i.e., her sensitivity to detect right and wrong features) was not as high as when she was finally able to experience emotions. An adequate level of virtue[96] is what will prevent emotions from misfiring, or rather, what will warrant a *correct* emotional acquaintance with moral act types and act tokens. I base this claim on what Aristotle has described as a virtuous emotional disposition:

> Both fear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way, is what is both intermediate and best, and this is characteristic of virtue. (*NE*, II, V: 1106b)

Charles Starkey has also relied on Aristotelian virtue to argue that emotions are essential to moral virtue (2008). He argues that emotions provide a *full understanding* of the situations that they involve. He says that, 'In such cases, emotions are not merely a symptom of the possession of an adequate understanding, but are rather necessary for having an adequate understanding.' (2008: 425)

Starkey argues that emotions provide a special state of awareness, based on a comparison with unemotional states. He argues that the *focus* and *import* of emotional states is unique given two assertions about the nature of emotion (Ibid: 430). The first assertion, according to Starkey, is that emotions include a cognitive and affective element (Ibid: 431). He takes the cognitive element to be a sort of "apprehension." He defines an apprehension as a 'phenomenal cognitive state that is distinct from an explicit belief about its object. It is like a perception of an object, if perception is understood in a robust way as a "seeing as," but the term apprehension is preferred because it avoids possible confusion in the use of the term perception.' (Ibid: 431) For example, an object of a given emotion can be either a dead animal *on sight*, or an imagined dead animal.

According to Starkey, the affective element of emotions includes bodily physiological changes as well as the previously mentioned psychic feelings (Ibid: 431). Starkey's second assertion about the nature of emotion is that an emotion is not just the occurrence of apprehension and feelings. He argues that a

[96] I am endorsing a view of disunity of virtue. In other words, I do not share the orthodox virtue ethics view that in order to possess one virtue one would need to possess all of them (See Srneevisan 2020: chapter 4).

certain apprehension and a certain feeling in an emotional state are mentally *associated* (Ibid: 432). For example, when Bob feels anger whilst Mary is insulting him, he usually makes sense of his rising heart rate in terms of Mary's insult rather than in terms of the noise of the birds singing outside.

As I have argued as well, Starkey claims that emotions can *focus* our attention through an *aroused* mental state, and generate awareness in this way (Ibid: 435). He argues that the awareness produced by emotions also allows for gathering more information. For example, when Mary is insulting Bob, there are several things that might be capable of grasping Bob's attention including other sounds, visible objects, etc. However, by being associated to a state of arousal (i.e., experience of anger), Starkey argues, Bob will notice Mary's insult rather than these other aspects of his environment, and by focusing on the insult he can gather more information, rather than simply acknowledging it and moving on to another object of his concern.

Additionally, there is an *import* of emotional episodes as described by Starkey, which is the 'experienced significance that a certain object or situation has for us.' (Ibid: 436) So, for example, awareness of rude behaviour might have a much greater import for Bob if he cares more about rude behaviour than for Mary, who does not. Starkey argues that 'emotions give states of awareness an import that they would lack otherwise, and emotional experiences are thus typically characterized by an experienced significance lacking in unemotional states.' (Ibid: 437) In other words, without emotion, Bob may pay attention to Mary's insult but not care about it so that it is not significant for him. On the other hand, if Mary's insult is associated with Bob's emotional reaction of anger, Bob will not be indifferent to the insult. Whilst experiencing anger, according to Starkey, Bob's apprehension of Mary's insult would 'seize him' and be of greater import (given that he already cares about rude behaviour). However, Starkey argues that it is not the case that the import of the object of apprehension is solely a product of emotion (when present). He argues that other beliefs and concerns may alter the sense of importance of a given object, although emotions:

> …play a crucial role in determining the initial import of the object presented to us. As in the case of focus, this greater import of the object of the emotion is accounted for by the *association* between the object of the emotion and the state of affective arousal. Here, the object of the emotion is made sense of, in part, in terms of the affective arousal, and the affective arousal is, in turn, made sense of in terms of the object of the emotion. (Ibid: 437, emphasis mine)

Starkey goes on to link the special awareness that emotions can provide to understanding. He argues that our understanding of an object depends to a great extent on our awareness of it (e.g., to be aware of a poisonous spider as a threat is typically to understand it as such) (Ibid: 439). Due to this relation, Starkey says, a change of awareness during an emotional episode typically affects the understanding of the object of the emotion. Recall that by focusing our attention, emotions will lead us to become more fully aware of the object or situation and such focus produces a richer understanding. Additionally, the

greater sense of import of the apprehended object that emotion provides can make us explicitly or tacitly understand it as *important to us* (Ibid: 439).

Finally, Starkey argues that to have *full understanding* involves an adequate understanding of an event 'given our larger values, interests, concerns, and goals (i.e., those things that relate to our thriving), and is a normative concept in that it implies a sufficiency or propriety of understanding.' (Ibid: 440) For example, Mary can be said to have full understanding of the dangerousness of spiders if, besides knowing that some spiders are poisonous and others are not she is also aware that her excessive fear of spiders—a concern of Mary's— can sometimes be misleading. In this case Mary has full understanding of the dangerousness of spiders given that she understands to what extent spiders are dangerous despite her fear.

However, although so far Starkey and I have argued that an adequate understanding of an object or a situation requires awareness and attention (see chapter IV)—and that emotional episodes can provide both of these requirements— there has not been an argument for the *necessity of emotions* to acquire understanding in general, nor for acquiring moral understanding. Starkey argues that there are 'several compelling reasons to think that it is a fact about human beings that emotion is necessary for many states of understanding.' (Ibid: 440) His claim is that emotions are necessary for many instances of full understanding, though not that emotions are necessary for the capacity alone to have full understanding. He does not argue either that emotions are always necessary for full understanding. He goes on to argue for three considerations[97] that show the necessity of emotions for acquiring many states of full understanding, although I will not discuss them here. My aim is to argue for the necessity of emotions for virtue and for the highest levels of moral understanding of the putative moral statues of act tokens and act types, not for the necessity of emotions for other tokens of full understanding. Shortly I will describe these levels, and defend why emotion is necessary for virtue.

MES intends to show the importance of the role that emotions can play in the acquisition of moral understanding, and so it stresses the ways in which emotions can be epistemically useful for at least gaining understanding of the character of others as well as our own, of situations, and of actions that we regard as *prima facie* moral. However, MES does argue that emotions are necessary for moral understanding when it comes to the following instance of moral understanding: *the highest level of moral understanding*. The way MES argues for this claim is by appealing to virtue. The way in which Starkey argues that emotion is necessary for virtue is by explaining that virtue is directly related to our thriving (Ibid: 444-445). And, given that emotions are necessary for possessing many instances of full understanding, one of which is full understanding of the values, interests, concerns, and goals that are conducive to our thriving, Starkey argues, emotion is necessary for virtue (Ibid: 445). However, Starkey

[97] They are (i) reflective intuitions, (ii) human information processing, and (iii) empirical evidence (Starkey 2008: 440-444).

does not defend a specific account of virtue. As explained at the beginning of this section, I adopt a McDowellian conception of virtue (as an ability to recognise what the right thing to do is in different situations) which draws on Aristotle (whose account I also deploy to ground the correct moral understanding and behaviour that results from virtuous emotion).[98] The link that I establish between emotion and virtue is based on the Aristotelian picture of the virtuous person as someone who emotionally responds 'at the right times, with reference to the right objects, towards the right people, with the right motive, and in the right way' (*NE*, II, V: 1106b). Due to these characteristics of the emotional reactions of virtuous individuals,[99] which are not misleading, excessive nor defective, every virtuous emotional acquaintance with *prima facie* moral situations and actions will be *correct*. This fact about the nature of virtuous character alongside the epistemic usefulness of emotion is what—according to MES—guarantees that virtuous emotional acquaintance with *prima facie* moral actions and situations constitutes the highest level of moral understanding.

A first-hand virtuous emotional experience is the first-hand emotional experience of a *virtuous person.* The *highest level* of moral understanding would consist at least in part in acquaintance or a first-hand virtuous emotional experience with *various* putative moral statuses of act types and act tokens. For example, Mary would possess the highest level of moral understanding of the wrongness of betrayal (act type), had she experienced many instances of betrayal (act tokens) via virtuous emotional episodes (e.g., betraying Bob and experiencing *guilt*, being betrayed by Bob and experiencing *sadness*, observing Jane being betrayed by Bob and experiencing *empathy*, etc.). In other words, the more instances of betrayal that Mary experiences accompanied by an emotional episode, the more she can learn about the ways in which betrayal is wrong or impermissible until it becomes the highest level of moral understanding of the wrongness or impermissibility of betrayal.

It is worth noting that according to MES, it is possible that correct emotional acquaintances can lead in part to the highest level of moral understanding of the putative moral statuses of certain actions, even if the agent that experiences them is not fully virtuous. In this sense, anyone who is capable of experiencing adequate emotions can possess the highest level of moral understanding of the putative moral statuses of various tokens of morally appraised types of actions.

It is now easier to see that according to MES, when Mary was inside the laboratory, she was not fully virtuous. First, as Aristotle has pointed out, the virtuous person is someone who has appropriate emotional responses (*NE,* II, V:1106b), and inside the laboratory Mary could not undergo emotional

[98] Another way to argue for a necessary link between virtue and emotion is Gopal Sreenivasan's (2020). Sreenivasan argues that emotions are a necessary component of some virtues. According to his Integral View, 'For some virtues, a morally rectified emotion trait is a functionally integrated constituent of the virtue.' (2020: 30) For example, a morally rectified *sympathy trait* is a functionally integrated constituent of the *virtue of compassion.* For the full account of the Integral View see Sreenivasan (2020).

experiences at all. Second, according to MES, the virtuous person is *also* someone who has acquired the highest level of moral understanding of the rightness and wrongness of several act types and act tokens through emotional experiences. Hence, due to her lack of emotional experiences inside the laboratory, there was no way in which Mary could acquire the highest level of moral understanding. In the next and final section of this chapter, I will briefly describe another account of emotional acquaintance in order to contrast it with mine.

iii.i Ballard's Account of Emotional Acquaintance

Brian Scott Ballard (2020) has pointed out that although many authors have argued that emotional experiences can provide some sort of epistemic benefit, so far, their proposals have failed to capture the distinctive way in which emotions epistemically enhance evaluative judgements (2020: 114). In particular, he argues that we can epistemically achieve everything that these views suggest through other means and not only through emotional experiences. Hence, he argues, these views cannot fully account for the distinctive epistemic benefit of emotions. He targets four influential views: **(i)** that emotions are required for the formation of evaluative concepts (see e.g., Prinz 2007; Goldie 2002; Vanello 2020), **(ii)** that emotions have important attentional effects (see e.g., de Sousa 1987; Deonna and Teroni 2012; Brady 2013), **(iii)** that emotions provide direct justification for evaluative beliefs because they are evaluative perceptions (see e.g., Roberts 198; Döring 2003; Prinz 2004; Tappolet 2016), and **(iv)** that emotions promote evaluative understanding (see e.g., Brady 2013; Deonna and Teroni 2012; Roberts 2013). His own view is that emotions *acquaint us with value*, and he goes on to argue that it is the most promising (Ibid: 114).

My MES is only committed to (ii) and (iv), rejects (i) (assumed by Moral Mary's thought experiment), and remains neutral to (iii) and Ballard's view. On the other hand, Ballard's *Acquaintance with Value* view explains that emotional experiences seem to 'confront' us with values. According to Ballard, emotional acquaintance provides *an instantiation* of value properties (2020: 121). In one of his examples, he explains that 'Dud' is someone who sees wild horses running *without* experiencing any emotion; Dud knows that the property of beauty is instantiated when he sees wild horses running before him, due to the fact that he may be acquainted with the non-evaluative properties that ground their beauty (such as the colours of their fur or the wavy effect that the wind has on their hair). However, Ballard explains that in this case Dud is not acquainted with beauty *itself*, or rather, he is not acquainted with the horses' features *qua* beauties (Ibid: 121). Ballard argues that even without direct perceptual acquaintance with the relevant non-evaluative properties, for example, by learning the badness of a situation via testimony, and just hearing the facts that determine the *badness* of something, Dud's emotional experience whilst hearing the facts would acquaint him with the badness of that something.

Ballard's view is explicitly realist. In other words, he holds a view that describes values to be in no way metaphysically dependent on emotion or any other mental state. His realism[100] in particular intends to show *how* realists should understand the relation between value and emotion. He argues that acquaintance with value is precisely the epistemic benefit that emotions can provide (Ibid: 113-114).

Here is the passage that illustrates emotional acquaintance with value according to Ballard:

> 'Now, I do not wish to speak here of degrees of knowledge. But in some sense the war veteran has—I am of course generalizing—a deeper grasp of the horrors of war. From what does this arise? Surely, his first-person experiences in battle. But I submit that, take the emotion out of his experience, and you take away that deeper grasp. It is not merely that the veteran observes for himself the shells bursting in sand. It is that, as he watches, he trembles, sickens, shudders, and whatever else. His emotional engagement with the horrors of battle is what facilitates his acquaintance with them as horrors. Without emotion, he may be acquainted with things that are horrors, but without emotion, he is not acquainted with their *horrificness*.' (Ibid:114)

There are three ways in which Ballard's view differs from MES. First, Ballard's view (henceforth AV) is explicitly realist about values. In other words, he argues that values (such as goodness and badness) are metaphysically independent from emotions or any other mental state, although at the same time he argues that emotion and value have a special epistemic connection (i.e., emotions acquaint us with value). In contrast, MES remains neutral about the metaphysical status of value, given that all MES requires is that certain actions are regarded as 'right' or 'wrong', 'permissible' or 'impermissible'. It does not require a theory that accounts for whatever makes certain actions as right or wrong, or that explains how such judgments are to be understood. Hence, many metaethical views can deploy it to explain some of the roles that emotional can play in the acquisition of understanding of [moral] value, thereby making MES as a more attractive and more explanatory view than AV.

Second, although AV suggests (through the war veteran example above) that first-hand emotional experiences acquaint us with a 'deeper grasp' of value, AV does not distinguish it from emotional acquaintance via testimony, for example. AV does not appear to have levels of acquaintance, which MES does. By pointing out that moral understanding (gained through emotional experiences) comes in degrees, MES explains to a greater extent many instances of our grasping of value.

[100] Ballard explains that his motivation for endorsing this kind of realism is justified, given that most of the time emotions play an essential role in anti-realist theories. His project aims to encourage realists to show that they also can satisfactorily account for the relation between emotion and value. (2020: 113)

Third, unlike AV, MES provides a more thorough explanation of what is involved in emotional acquaintance with putative values by providing a detailed account of the components of emotion, and by relying on a conception of virtue.

Hence, MES vindicates the epistemic significance of emotional experience in a more thorough way than AV.

In the next chapter, I will elucidate the last argument in favour of the claim that emotions are useful for acquiring moral understanding, which is that emotional episodes can change our moral perspectives. My aim is to show that emotions can both increase and *improve* our moral understanding.

Conclusion

In this chapter I argued that the highest level of moral understanding consists in part in an adequate *first-hand emotional experience* with tokens of morally appraised types of actions. A first-hand *virtuous emotional experience* is the emotional acquaintance with putative moral actions of a *virtuous person.* However, according to MES, if an agent who is not very virtuous undergoes various appropriate first-hand emotional experiences can come to acquire the highest level of moral understanding of the putative moral statues of certain actions (e.g., the wrongness of betrayal).

I also argued that moral understanding comes in degrees (i.e., low, medium, high, highest), and described three ways in which emotional episodes can provide us with upstream, downstream (and both) instances of moral understanding.

Finally, I compared Ballard's account of emotional acquaintance with mine. I argued that MES is a more attractive and more explanatory view than Ballard's realistic account of AV for the following reasons: (i) MES remains neutral about the metaphysical status of value and thus can be deployed by many metaethical views, (ii) AV does not include levels of acquaintance and therefore does not explain as many instances of understanding as MES, and (iii) Unlike AV, MES provides a detailed account of the components of emotion.

Chapter VI

Can Emotional Episodes Change Our Moral Perspectives?

My aim in this chapter, is to argue that a certain kind of emotional episode can change our moral perspective, and that this change can be positive (but not invariably so). By 'moral perspective', I am referring to *the epistemic standpoint whereby agents identify the morally relevant features that ground the set of general moral beliefs that they implicitly endorse* (e.g., 'acts of charity are right', 'lying to your friends is wrong'). I argue that, since typically moral emotions (e.g., compassion and guilt) tend to make certain features of moral actions striking or *salient*, their experiencing them can direct our attention to new or different morally relevant features, thereby producing a change in moral perspectives, which in turn can involve an improvement in our moral behaviour.

First, I briefly introduce some views that identifies certain emotions as *moral* (Gibbard 1990) (Haidt 2003) (Prinz and Nichols 2010). I argue that an experience of *typically moral emotions* can involve a change in moral perspectives, given that these emotional episodes can constitute, arise from, or are associated with moral judgements (Prinz and Nichols 2010:112). Second, through some examples, I argue that a shift in our second-order moral views —i.e., those views and/or attitudes about the appropriate grounds of moral judgements (Sinclair 2021:194)—brought about by an emotional episode can lead to a change in moral perspectives. For example, Mary's episode of *anger* when John is bullying her friend Bob can lead her to disapprove of that form of violence, and her disapproval can modify her views on the morality of bullying (i.e., she forms the first-order view that bullying is wrong). Hence, Mary's new second-order view (i.e., that one's moral judgements should be sensitive to the presence of violence in the object of evaluation) would ground as appropriate the moral view of regarding bullying as wrong. In this sense, her episode of anger can be regarded as *moral*. In other words, emotional episodes can be regarded as moral *in the sense that they can change our moral perspectives.*[101]

Finally, I argue that this change of moral perspectives can involve a change in what an agent usually recognises to be the morally relevant features of moral actions, and therefore can result in an improvement of moral understanding. This improvement would consist in the coherence between two abilities: (i) the ability to *recognise* new morally relevant features of moral actions, (ii) the ability to *react appropriately* (i.e., to experience adequate moral emotions and to perform right actions). For example, suppose that, inside the laboratory, Mary used to regard racist jokes as harmless. After an incident where she became the victim of racist jokes outside the laboratory, Mary experiences *resentment*. Resentment focused her attention on the tone with which she was spoken to, as well as on

[101] In contrast, for instance, Ben-Ze'ev claims that emotions can be called moral if (i) Their core evaluative concern (or appraisal) is moral. For example, Mary's guilt evaluates her action as wrong, so this makes it a moral emotion. And (ii) emotions lead to beneficial moral consequences. For example, Mary's compassion motivates her to help others.

other factors, like the way it feels to be singled out because of one's race, causing her to disapprove of it. Consequently, her perspective on the morality of racist jokes changes, and now she does not hold those jokes to be morally harmless, but *prima facie* wrong. Her moral understanding of the wrongness of racist jokes will have improved if (i) acts of discrimination are now salient to her (ii) she experiences indignation, sadness, resentment, anger, regret, or any negative emotion associated with making fun of other people's backgrounds, and if she refrains from making racist jokes. I argue that if Bob had been the victim of a racist joke, and Mary had become emotionally acquainted with this fact by also experiencing resentment, the effect would have been *almost* the same. As argued before, a first-hand emotional acquaintance can provide the highest level of moral understanding. Emotional acquaintance via testimony or through the experience of somebody else can allow us to gain some moral understanding, although not as much as when emotional acquaintance is first-hand, as it was discussed in the previous chapter.

Therefore, supposing again that inside the laboratory Mary used to judge that racist jokes were morally harmless, and outside the laboratory she witnesses Bob being the victim of racist jokes (or is told so by him), and she experiences resentment, her resentment might focus her attention on the tone Bob was spoken to, as well as on other factors. In this case, Mary's perspective on the morality of racist jokes also changes, and she comes to judge that racist jokes are wrong. Suppose as well that acts of discrimination are now salient to her, and that she experiences negative emotions associated with making fun of other people's backgrounds, and she refrains from making racist jokes in the future. Mary's moral understanding of the wrongness of racist jokes also improved, but nonetheless not *as much* as it would have had she being the victim of the racist jokes herself. In other words, by experiencing being the victim of racist jokes herself and by experiencing resentment, Mary can come to understand to a higher (or to the highest extent) the wrongness of racist jokes, given the *stronger salience* of the features that make racist jokes wrong (such as the tone in which they are told).

i. Typically Moral Emotions

It seems uncontroversial to claim that emotions are primarily social, since they involve action tendencies that affect our interactions with one another (e.g., *admiration* for Bob can prompt Mary to *praise* him and to *imitate* him) (Gibbard 1990:138).

However, it is possible to classify some emotions as those which relate to the coordination of our moral practice, or as those which at least play a role in regulating our behaviour so that it conforms to established prospective *norms* (i.e., to what is believed one ought to do and ought not to do). For example, Mary feels *indignation* towards bullfighting (i.e., she morally disapproves of the practice), and so regards it as something that ought not to be done, since as part of our moral behaviour, inflicting

suffering on other sentient beings for leisure should be avoided. In the same line, Jonathan Haidt has famously defined moral emotions as ‘‘those emotions that are linked to the interests or welfare either of society as a whole or at least of persons other than the judge or agent.’’ (Haidt 2003:853)[102] The emotions that Haidt lists as prototypically moral are *compassion*, *anger*, *elevation*, and *disgust* (Ibid: 854). In contrast, other emotions such as *frustration* involve different action tendencies that do not relate to the coordination of our moral practice. Frustration is usually not caused by disapproval of one's or others' action. If it is, in general the perceived transgression is not of an established norm associated to moral behaviour. Rather, frustration tends to be caused by unfulfilled expectations. For example, Mary can feel frustration towards Bob for not remembering her birthday. Mary feels frustrated due to the fact that Bob is not meeting her expectations of congratulating her, and maybe buying a present for her. Even if Mary regards Bob's action as something that he ought not to have done, forgetting the birthday of one's friends is generally not considered to be a transgression of an established moral norm.

It also seems possible, as Gibbard notes, to describe emotions that concern ethics as distinct 'adaptive syndromes' with typical causes, typical expressions, and typical action tendencies (Ibid., 136). *Guilt*, for example, is often caused by having wronged somebody else, it involves a consciousness of having done wrong, and it conduces to amends. *Shame,* on the other hand, emerges when one judges that there is something wrong with oneself, and that this might be a helpless situation: 'One feels guilty for having told a lie, but one feels shame for being a liar. Guilt concerns transgressions; shame involves shortcomings. Guilt urges reparative action; shame encourages social withdrawal.' (Joyce 2000: 102)

Hume has famously spoken of moral *sentiments,* which would arise from the psychological mechanism of *sympathy,* that explains how we come to feel what others are feeling (T 3.3.1/575). Moral sentiments would be "feelings of approval or disapproval, praise or blame, esteem or contempt" (T.3.3.1). These sentiments would be the source of our moral ideas of goodness and badness. Given that both *sentiments* and *emotions* belong to the same Humean category of perception (i.e., they are both *impressions*), I am hereby taking the liberty to treat them as equal terms. For Hume, then, moral emotions can be distinguished from other emotions by describing their function of either expressing general approval, or expressing general disapproval (T3.3.1/581).

Now, it is widely accepted that one of the roles that emotions play, is that they are sources of *motivation*. Even those who think that reasons alone can motivate, they are also likely to admit that when present, emotions can motivate (Prinz and Nichols 2010:112). In a moral context, emotions typically motivate as a consequence of a moral judgement, although they can also motivate in the absence of one (e.g., *pride* can motivate Mary to become altruistic in order to gain popularity among her friends, and not because she believes it to be the morally right thing to do). Authors like Prinz and Nichols (2010), Ben-

[102] In contrast to what Haidt's view emphasises, moral emotions also seem to be concerned with oneself (e.g., it is possible for Mary to feel guilty in a situation where she did not stand for the moral principles that she believes in, since she considers that this is an important aspect for her personal moral life.)

Ze'ev (2000), and Haidt (2003), have identified three different families of emotions that arise in paradigmatic contexts where moral rules have been established (e.g., stealing, giving to charity). The first family would be the 'prosocial' emotions, which are *other-regarding* and promote morally good behaviour: *empathy, sympathy, concern, and compassion* (Prinz and Nichols 2010:122). The second family would be the *self-blame* emotions: guilt and shame. Lastly, the third family would be comprised of the *other-blame* emotions: *contempt, anger, and disgust*.

The fact that typically moral emotions are divided according to their objects is important for the purposes of this essay, since the evaluation they involve depends on whatever they make salient (i.e., on whatever they direct our attention to). There are other emotions that have been considered to be morally significant, which may serve the function of rewarding for good behaviour, such as *gratitude* and *admiration*, and others which may promote morally relevant commitments, such as *loyalty* and *love* (Ibid, 122-123).[103]

So far in this section, I have only briefly mentioned some aspects of what some thinkers have described as typically moral emotions. And so moral emotions can be: distinct adaptive syndromes, the source of our moral ideas, sources of motivation when associated to moral judgements, the other-regarding emotions (i.e., prosocial), and generally those emotions that arise in situations where moral norms are at stake, and which can play a role in coordinating our moral practice (e.g., by expressing our approval or disapproval towards certain actions). By 'coordinating' or 'regulating' our moral behaviour, this distinctive functionality of emotions would also involve correcting certain attitudes and actions (e.g., condemning mistreating animals for fun). The case of *guilt*, which is a 'self-directed' emotion (Joyce 2007:101), shows that when one judges to have violated some moral norm (e.g., revealing someone else's secret to their enemies), an amend must be made, and provided that this amend is carried out, one can be said to morally improve. Therefore, moral emotions can be interpersonally corrective, as well as self-corrective.

In this chapter, the two most important features of typically moral emotions, are (i) that they can be self-corrective and therefore lead to moral self-improvement, and (ii) the fact that they can constitute, arise from, or be associated with *moral judgements*. When I describe my view on the change of moral perspectives in section iii., the emotions that partially explain these perspectives are associated with our moral judgements. In my view, it is in this respect that any emotion that is associated with moral judgements can be categorised as 'moral.' Given that the evaluative view of emotions that I endorse takes emotions to at least six components (intentionality, cognition, personal evaluation, disposition or readiness to act, i.e., motivation, feeling, dynamism, see section i., page 7), it will be possible to describe

[103] It seems possible to derive some emotions from others (or to describe some of them in relation to others). For example, Bob's *joy* for Mary's promotion can be understood as a type of *sympathy*. Gibbard describes *resentment* and *indignation* as a type of *anger* (Gibbard 1990:126).

in which way they involve self-improvement. Additionally, my view is compatible with the claim that different moral emotions can be distinguished by their patterns of salience, and this feature (i.e., that they involve patterns of salience), is essential to describe how these emotions can lead to a change in moral perspectives, and self-improvement.

In the next subsection, I will describe the link between typically moral emotions, and their role in fixing patterns of salience.

ii.i What Aspects of Moral Action Are Made Salient by Typically Moral Emotions?

In the first and second sections of this chapter, I have pointed out the fact that every emotion seems to have a different object of evaluation. For example, *anger* might highlight features of the actions of those who have committed an offence (to others or to ourselves), *guilt* focuses on the aspects of our own actions. Roughly speaking, each emotion directs our attention to specific things (e.g., *grief* focuses on loss, *empathy* on the feelings of others), and therefore each emotion makes different things noteworthy (i.e., every emotion involves different norms or patterns of salience) (Gibbard 1990; de Sousa 1987; Elgin 2007; Ben-Ze'ev 2010).

And, just like the *fear* Mary experiences when she finds herself before a pack of wolves in the middle of the woods focuses her attention solely on the wolves (making their movements salient), fear also prompts Mary to elaborate an escape strategy (i.e., it motivates or causes her to act in a certain way). Similarly, the *guilt* she experiences after having betrayed Bob, focuses her attention on the lies she made him believe, and motivates her to confess or make amends. Therefore, it seems that the patterns of salience caused by emotions relate to their action tendencies[104] (e.g., guilt motivates compensation, shame leads to withdrawal, and contempt leads to aversion) (Prinz and Nichols 2010:135; Mason 2003: 241).

It then seems that depending on their object of evaluation, emotions will make different things salient, and given their specific action tendencies, they can be understood to have different functions. In chapter IV, I made an analogy between emotions and desires, relying on Schroeder's example of Directed-Attention Desires (Schroeder 2007:147). The analogy is meant to show that, just like certain emotions will make certain things salient and lead us to act in a specific way (e.g., guilt focuses our attention on

[104] For example, according to Gibbard, 'the concept of guilt is given by a cluster of circumstances in which it will normally be felt, by a set of normal expressions, and by tendencies to action it normally stirs.' (1990:150).

our offence towards someone else and prompts us to make amends), certain desires will make certain things salient, and these things in turn will prompt us to act in a certain way (e.g., desire for a cup of coffee focuses our attention on the coffee shops on our way to work, and it prompts us to walk towards one and buy coffee).

Hence, I argue that emotions which can be regarded as *moral* usually can make salient—or direct our attention to— the *topics* and *considerations* that promote certain moral actions (i.e., actions that should be done), depending on their object of evaluation. As suggested at the beginning of this subsection, moral emotions will make salient different features.

So, generally:

a) Given the perceived features *WZ* of action *S*, moral emotion *X* prompts to action *Y*. And: Moral emotion *X* is associated with a moral judgement *F* justified by reference to features *WZ*.

For example:

Mary and Bob perceive betrayal and mockery (*WZ*) of Mary's act of bullying towards Bob (*S*). Mary experiences guilt (X), and it prompts her to apologise to Bob (*Y*).

And: Guilt (*X*) is associated with the moral judgement "I have done something wrong because I have betrayed and mocked Bob." (*F*)

In other words, emotion can prompt us to take notice of features that justify the moral judgement.

b) Moral emotion *X* directs attention to features *WZ* of action *S*, and prompts to action *Y*.[105] And: Moral emotion *X* is associated with a moral judgement *F*.

For example:

Guilt (*X*) directs Mary's attention to the betrayal and mockery (*WZ*) of her act of bullying towards Bob (*S*), and guilt prompts her to apologise (*Y*).

And: Guilt (*X*) is associated with the moral judgement "I have done something wrong because I have betrayed and mocked Bob." (*F*)

[105] Moral emotion X can always direct attention to several features of a certain action, and prompt to perform more than one action.

c) Moral emotion *X* prompts to action *Y* and leads to the perception of features *WZ* of action *S*. And: Moral emotion *X* is associated with a moral judgement *F*.

For example:

Guilt (*X*) prompts Mary to apologise (*Y*) and leads her to perceive betrayal, mockery, Bob's sadness and disappointment (*WZ*) of her act of bullying (*S*).

And: Guilt (*X*) is associated with the moral judgement "I have done something wrong because I have betrayed and mocked Bob."(*F*)

In all of these cases, it is possible to experience more than one moral emotion at once (e.g., (*guilt* and *compassion*). Also, it is possible that the same moral emotion is associated with more than one moral judgement (e.g., guilt is associated with the moral judgement "I have done something wrong" and "It is not right to mock one's friends." In the next subsection, I will discuss some ways in which emotional episodes can lead to a decrease in moral understanding.

ii.ii Decrease in Moral Understanding

As discussed before in chapters III and V, emotional episodes can easily lead to a decrease in moral understanding or hinder us making true moral judgements. Let's go back to the example in section i., where Mary is angry at Bob for having forgotten her birthday. Remember that Mary (as all of us) is disposed to 'see' or perceive the world a certain way. More precisely, Mary considers certain features of the world to be worthy of notice (i.e., she cares about them), such as the actions performed by her family and friends, the weather, animals, political news, etc. In a loose sense then, it is fair to say that she is disposed to notice what her friend Bob does. In a stronger or more particular sense, it is also possible to speculate both that birthdays are very important for Mary, and that she is primed to be in search of things other people do when there is a day which especially concerns her. It was also assumed that a condition for possessing certain emotions, is that one is implicitly disposed to feel them, and in this case, Mary is someone who is capable of feeling anger. To be clear then, emotional episodes that make certain things salient and that can lead to certain actions, at least depend on (i) that the agent is disposed to notice features *WZ*, and (ii) that the agent is disposed to feel emotion *X*.

Therefore, once Mary notices[106] that Bob forgot to congratulate her for her birthday, she gets angry, and her anger, according to the patterns of salience view, directs her attention to other features of Bob's action (i.e., forgetting). For example, that he did not buy her any presents, nor a cake. In light of these features made salient by anger, Mary's anger also prompts her to yell at Bob, and she implicitly or explicitly judges 'Bob has wronged me.' This case is not yet a clear example of how moral episodes can go wrong, but it is easy to imagine that in this situation, if Mary's anger is disproportionate, it might direct her attention to other features of Bob's action or of Bob himself which are not related (or relevant) to what she has perceived as a fault on his part (such as the lack of repentance of the tone of his voice when he is trying to apologise, or the fact that he is not dressed fancily for her special day). Consequently, these ongoing perceptions of irrelevant features made salient by anger can cause Mary to stop talking to Bob for years after this incident, an action which can be characterised as inappropriate or disproportionate, given the nature of the offence.

A similar misfiring situation can happen in the case of a moral emotion. For example, let us assume that Mary is also both disposed to notice whenever some people harm others (such as instances of bullying), and to feel *compassion*. She hears about John bullying Bob, and she judges (implicitly or explicitly) that Bob has been wronged. Mary's compassion might make salient the fact that Bob is crying, and that he does not have many friends, and this prompts her to comfort Bob, and to be an especially kind friend to him. There seem to be at least four ways in which Mary can experience compassion inadequately.

First, compassion can make striking irrelevant features related to what has happened to Bob, such as the fact that the name of Bob's bully is John (like her grandfather's), and this can distract her from comforting Bob, thereby failing to be a good friend. Second, compassion can make striking subsequent relevant facts, such as the fact that Bob tends to be bullied by people who he has bullied in the past. However, despite this realisation, instead of pointing this out to Bob in order to make him reflect on his own moral behaviour, this instance of compassion motivates her to remain silent and support Bob in his suffering, thereby failing to do the right thing. Third, Mary's compassion identifies Bob's general vulnerability, and it prompts her to manipulate him. Consequently, compassion would fail as a moral emotion that is believed to prompt to aid. Finally, in fourth way, Mary's compassion can be disproportionate and fail as a moral emotion whereby by making Bob's vulnerability salient, it motivates her to be overly protective of Bob, which then incapacitates him to defend himself.

Hence, in general, moral emotional episodes can misfire in four ways: (i) by directing our attention to features irrelevant to the moral evaluation of the case (ii) by silencing *prima facie* evident relevant considerations, (iii) by obscuring morally relevant features of actions and situations, and (iv) by not prompting us to act in the way that we generally believe they should. In the next section, I aim to provide

[106] Mind that Mary could get angry and undergo the same psychological experience, either by imagining or remembering that Bob forgot about her birthday.

an account of moral perspectives, and of how adequate experiences of morally emotional episodes can lead to a change in moral perspectives for the better. I will also describe the kinds of moral understanding that these emotional episodes create after transforming these perspectives.

iii. Change in Moral Perspectives

Roughly speaking, moral perspectives are a set of views about which features of situations are relevant to the moral evaluation of those situations. What we are disposed to identify as morally salient features within a given moral action or situation, will also influence the formation of our moral judgements, and these in turn will influence our action tendencies. Moral emotions and moral perspectives are closely related, since this kind of emotions can direct our attention to the salient features that prompt to moral action.

Moral perspectives would work as a kind of moral lens. At the beginning of this chapter, I defined a moral perspective as *the epistemic standpoint whereby agents identify the morally relevant features that ground the set of general moral beliefs that they implicitly endorse*. Call the set of general moral beliefs 'first-order' (Sinclair 2021:195). This epistemic standpoint however, is at least partially formed and regulated by another set of views which are taken to be 'second-order' (Ibid: 194). In the case of moral perspectives, second-order views would evaluate whether the formation of our general disposition to find certain things as salient is *appropriate*. For example, Mary's set of second-order views accepts as morally appropriate to be disposed to identify as salient the needs of others, whereas the disposition to identify physical appearance as a morally relevant would be inappropriate.

Given that the formation of a certain moral perspective (MP)—for example, to disapprove of acts of discrimination, given the disposition to identify racist attitudes as salient— depends at least in part on the agent's second-order views, a change in MP would involve *a shift of focus* on what we regard as salient (and therefore as morally appropriate). I argue that this shift of focus can be caused by a certain kind of emotional episode.

For example:

Love in the famous example of Iris Murdoch's *M* and *D*. M is D's mother-in-law, and she dislikes different aspects of D's personality. She is also annoyed at the way she speaks, and at the fact that she behaves too informally. An interesting feature of the story though, is that M's behaviour towards D is nice and polite. However, a time comes when M's *love* for her son fixes her attention in other aspects of D's character. Suddenly, what appeared to be contemptible or impolite, now seems refreshing and unique. D's accent, for example, is no longer relevant in M's assessment on how fitting as a partner D

is to her son. What seems fitting to M now, is that D is friendly and generous (see Murdoch 1970:17-18).

Following the general theory of moral emotions in section ii.i, in this case, given that love for her son caused M to notice other features of D's behaviour:

b) M's love (*X*) directs her attention to friendliness and generosity (*WZ*) found in the way D speaks and behaves (*S*), and prompts her to perform genuine acts of kindness towards D (*Y*).
And, love is associated with the moral judgement 'D is a fit partner for my son because she is friendly and generous.' (*F*)

It might seem that *love* in this example is not a moral emotion in the strong sense, given that the judgement that it is associated with is not a judgement about 'right' , 'wrong', etc. However, Murdoch's case of M and D is an excellent example of a change of moral perspective caused by an emotion, given that Murdoch also treats the example in the context of discussion about moral behaviour. Murdoch argues that before experiencing love for her son changed the way she perceived D, M's fake acts of kindness towards D were not morally worthy[107] (Ibid: 18).

Put simply, M used to think that considerations such as the accent of a person and the fact that they behave informally were relevant to the moral appraisal of D. But M now thinks that friendliness and generosity are relevant to the moral appraisal of D. It seems that her change in perspective is shown by: (i) the shift of her attention on *new* features about D's character, (ii), the change in her judgement about D, and (iii) the shift of attention from irrelevant features such as D's accent and her lack of fancy manners to the fact that she is friendly and generous. This last shift in her second-order moral views about what she finds now appropriate, is explained by the fact that she does not identify D's accent and informal behaviour as negatively morally salient any longer. In turn, her first-order view about a fit spouse for her son changed from something such as "A fit spouse for my son is an overly fancy spouse" to "A fit is spouse for my son is someone friendly and generous."

In the next subsection, I will argue that another feature of a change in moral perspectives caused by emotional episodes is that it can lead to improvement in moral understanding.

iii.i Change in Moral Perspectives and Improvement in Moral Understanding

A possible outcome of Mary's new emotional experiences on leaving the lab was that they misfired. In chapter V I argued that that was the case when Mary's love for her bullfighter father led her to approve

[107] This case illustrates how emotional episodes can also lead to morally worthy actions. However, discussing this role of emotions is outside of the scope of this thesis.

of bullfighting. Then, Mary's moral perspective on bullfight changed when she stopped judging that bullfighting was wrong, but it is still unclear why her judging bullfighting is right did not improve her moral understanding.

Mary's father has reasons to believe that bullfighting is appropriate, and therefore believes that it is morally right. Mary's father believes that there are thirty encyclopaedia tomes that explain why bullfighting is an art. According to him, bullfighting is a ritual rather than a mere torture practice, and the clothes and the music involved in it are beautiful. He thinks that bullfighting connoisseurs know more about the preservation of bulls than most people, and even love and admire bulls more than most people. He thinks the bullfighter's skill and bravery is worth of great admiration: after all, the bullfighter is willing to die for his profession by putting his life at risk. If bullfighting is cancelled Mary's father says, the bull species might become endangered; bullfighting has become an iconic form of Spanish culture and art, and so it creates a sense of national pride among Hispanic countries.

The further criteria that I am introducing to judge whether Mary's moral perspective on bullfighting was correct or incorrect, relates to whether her moral understanding has improved in general. According to MES, the perfecting of the following abilities will determine whether her moral understanding has improved:

(i) the ability to *recognise* the morally relevant features of moral actions.
(ii) the ability to *react appropriately* (i.e., to experience adequate moral emotions and to perform the right moral actions).

Hence, an indication of moral improvement would consist in the *acquisition* and *reinforcement* of these abilities. If Mary's love for her father leads her to judge that bullfighting is right, it is possible to say that her judgement is incorrect and that love for her father is misfiring if it goes against the common point of view that inflicting suffering in sentient beings is wrong. Also, it is possible to say that her judgement is incorrect if she did not become more efficient at abilities (i) and (ii), showing that her moral understanding has not improved.

Consider how the following literary example illustrates that a morally emotional episode can involve a change in moral perspectives, and that it can result in an improvement in moral understanding. When accurate (i.e., when it corresponds to the common view of what is right and wrong, and therefore true) the change in moral perspectives would be *corrective*.

Consider the following excerpt from the novel *The Scarlet Letter.*

Emotion: *admiration*

> Individuals in private life, meanwhile, had quite forgiven Hester Prynne for her frailty; nay, more, they had begun to look upon the scarlet letter as the token, not of that one sin, for which

she had borne so long and dreary a penance, but of her many good deeds since. "Do you see that woman with the embroidered badge?" they would say to strangers. "It is our Hester—the town's own Hester, who is so kind to the poor, so helpful to the sick, so comfortable to the afflicted!"

Hawthorne, Nathaniel (1878):171. *The Scarlet Letter*, Lerner Publishing Group, 2017.

The Scarlet Letter is a story about a woman (Hester Prynne), who is punished by the members of her puritan community for having committed adultery. Her punishment consists in wearing an embroidered letter 'A' in all of her clothes, and in being rejected by everyone who lives in the same village. The villagers view Hester as a sinner deserving of shame and rejection, and so they speak badly of her and somehow insult her whenever they see her. However, after witnessing Hester's loving behaviour towards those in the community who suffer, the villagers' change of moral perspective on what Hester did before (i.e., committing adultery) can be explained by the general theory of moral emotions:

(a) Given the perceived features *WZ* of action *S*, moral emotion *X* prompts to action *Y*. And: Moral emotion *X* is associated with a moral judgement *F*.

For example, Hester performs comforts the suffering (as perceived by the villagers), and they experience *admiration*.

Given the perceived kindness and helpfulness (*WZ*) produced by Hester's aid to the poor, the sick, and the afflicted (*S*), admiration (*X*) prompts to forgiveness (*Y*).

Admiration in this case can be associated with the moral judgement 'Hester Prynne is a morally good person because she is kind and helpful to everyone.' (*F)*

Admiration shifted the villagers' attention from focusing on Hester's adultery, to focusing on her good deeds.

It is possible to say that the villagers' moral understanding improved if:

(i) They have acquired the ability to *recognise* the morally relevant features of Hester's moral actions. Even if her adultery was salient to them given their religious beliefs, admiration is prompting them to regard Hester's adultery as less relevant or less salient.

(ii) They have acquired the ability to *react appropriately* towards Hester Prynne (e.g., if they experience sympathy, compassion, or in general positive emotions towards her).

In this case, as opposed to the case of Mary's and her beliefs about bullfighting, admiration is a moral emotion in the sense that it is correcting the villagers' moral behaviour towards Hester. Outside the supposed Puritan village where Hester lives, the common point of view of other societies would probably not shun Hester (or at least not in the extremely uncharitable way in which those villagers

did), even if these societies disapproved of adultery. The shift in the villagers' second-order moral views about what are the appropriate features on which to base a judgement, is explained by the fact that they do not identify Hester's past mistake as negatively morally salient any longer. In turn, their first-order view about finding Hester worthy of shunning changed from something such as "An adulterous person should be shunned" to "Anyone can make mistakes and that does not stop them from being good."

Conclusion

I have attempted to show that certain emotional episodes can change our moral perspectives. Although there are some ways in which they can misfire, typically moral emotions can shift our attention to morally relevant features that we did not previously consider relevant, thereby changing our second-order views. This change in our second-order views can lead to a change in our first-order views, which in turn can lead to an improvement in our moral understanding.

Conclusion

In this thesis I argued that emotions can play a distinctive role in the acquisition of moral understanding. I based my argument on two features of emotions: (i) their components (i.e., intentionality, cognition through patterns of salience, evaluation, motivation, and feeling) and (ii) the fact that emotions can direct our attention to the morally relevant features of those actions which we judge to be either right or wrong. I also argued that certain emotional episodes can lead to the acquisition of the highest level of moral understanding, and that they can lead to a change in our moral perspectives in a way that improves our moral understanding. I argued for these claims in the following way.

In the first chapter, I first introduced the hypothetical thought experiment of *Moral Mary*, who has spent her whole life inside a moral laboratory emotionally sedated. I explained that inside the laboratory, Mary acquired moral knowledge and moral understanding to a great extent. Second, I endorsed Simple-K Reliabilism as an analysis of knowledge, given that it is compatible with various metaethical views, and that it is more compatible with an Inclusive Project in epistemology. Third, I argued for the possibility of moral knowledge, and how different metaethical views can adopt it. Fourth, I compared Hills' (2009) and Sliwa's (2017) accounts of moral understanding. Whilst Hills' account of moral reasoning abilities illustrates one way in which moral understanding can be acquired, exercised, and measured, Hills' account does not include emotions as an important mental state for moral understanding and cannot explain the highest level of moral understanding as explained by my Moral Epistemic Sentimentalism (MES). I introduced the claims that ground MES, and argued that it is a view that emphasises the importance of the role that emotions can play in the acquisition of moral understanding. In contrast, Sliwa's Moral Knowledge account does regard emotional experiences as highly important for the acquisition of moral understanding. Nonetheless, it does not provide an explanation of the components of emotions and other ways in which emotions play a crucial role in the acquisition of moral understanding, which (MES) aims to do.

In the second chapter, I argued that moral deference is problematic due to an epistemic deficiency: it cannot provide moral understanding of the highest level. I first discussed moral pessimism (Hills 2009; McGrath 2011), and then Fletcher's psychological pessimism. Second, I described Wodak's optimism about moral deference (2019), and replied to the two previous kinds of pessimism. I then discussed Callahan's moral pessimism (2018), and argued that although she includes emotions in her picture of moral understanding, her pessimism about moral deference does not describe the most relevant deficiency involved in the acquisition of moral understanding via moral deference. Finally, I described my Moral Epistemic Pessimism about moral deference, and argued that moral deference is problematic given that it cannot provide the highest level of moral understanding, due to the receiver's lack of a first-hand emotional experience with the morally relevant features of a certain action.

In the third chapter, first I roughly described different theories of emotion. I categorised my view on emotion as part of the evaluative category, and I argued that emotions are complex mental states with six components. I then described three ways in which some authors have argued that emotions can be epistemically helpful, but also some ways in which emotions can be epistemically unhelpful. Following Elgin (2008), I pointed out that although emotions can be variable and volatile, they can also be refined. Third, I differentiated epistemic emotions from moral emotions by arguing that moral can constitute, arise from, or be associated with moral judgements. I also discussed the Perceptual Theory of emotions and distinguished it from MES. Finally, I discussed Goldie's view that some emotions can be systematically misleading in order to argue in favour of a healthy skepticism towards the putative positive epistemic roles of emotions.

In the fourth chapter, I argued that exercising attention is a necessary condition to acquire understanding. I argued that, although emotional episodes are not necessary for the acquisition of understanding in general, they can focus our attention in a special way. I briefly, compared emotions to desires, and argued that the necessary feeling component of emotions makes them better candidates for the acquisition of understanding.

In the fifth chapter, I argued that the highest level of moral understanding involves in part a correct first-hand emotional experience with tokens of morally appraised types of actions. I also argued that moral understanding comes in degrees, and described the ways in which emotional episodes can provide us with different instances of moral understanding. I explained that once Mary left the moral laboratory and was able to experience emotions, her moral understanding increased. According to MES, the phenomenology and direction of attention that some emotional episodes can provide, lead to the acquisition of moral understanding in a distinctive way. MES is a democratic view of acquisition of moral understanding (as opposed to elitist), in the sense that an agent who is not very virtuous though undergoes various appropriate first-hand emotional experiences can come to acquire the highest level of moral understanding of the putative moral statues of certain actions (e.g., the wrongness of betrayal).

In the sixth chapter, I argued that certain emotional episodes can lead to a positive change in moral perspectives. I defined a moral perspective as *the epistemic standpoint whereby agents identify the morally relevant features that ground the set of general moral beliefs that they implicitly endorse.* What we are disposed to identify as morally salient features within a given moral action or situation, will also influence the formation of our moral judgements, and these in turn will influence our action tendencies. I provided three kinds of schema which describe some ways in which a change in moral perspective can occur via an emotional episode.

MES succeeds as an account that emphasises the importance of emotional episodes for moral understanding. It provides a description of the components of emotions that make them suitable for the evaluation of putative moral actions, and relies on the epistemic faculty of attention (on which the

research is extensive). MES provides three novel claims regarding the acquisition of moral understanding: (i) emotions are necessary for the highest level of moral understanding, (ii) emotional episodes can change our moral perspectives thereby improving our moral understanding, and (iii) acquisition of moral understanding via moral deference is epistemically defective given that it does not lead to the highest level of moral understanding.

Another distinctive feature of MES is that it elucidates levels of moral understanding, which few theories of moral knowledge do. MES presents emotions in a very positive light, although it is also aware of the many ways in which emotions can misfire.

On the other hand, MES has yet to come up with answers or explanations as to how certain emotional episodes *undermine* understanding. Although the fact that emotions focus our attention can lead to an increase in understanding, by consuming our attention emotions can also lead us astray and not direct our attention towards what is relevant in a given situation. MES still needs to expand its research on the nature of moral emotions, and to further argue in which ways it is a democratic view of moral understanding. In particular, MES aims to continue to develop the claim that involuntary cases of emotional episodes can lead to an improvement in moral understanding.

References

Alston, W. P. (1967) ‘Emotion and feeling’ in *The Encyclopaedia of Philosophy* New York: Macmillan

Aristotle. *Nicomachean Ethics*. Translated by S. Broadie (2002) Oxford: OUP

Armon-Jones, C. (1985) ‘Prescription, explication and the social construction of emotion’ in *Journal for the Theory of Social Behaviour, 15*(1), 1–22

Armstrong, D.M. (1973) *Belief Truth and Knowledge* London: CUP

Arnold, M. (1960) *Emotion and Personality: Psychological Aspects* New York: Columbia University Press

Audi, R. (2019) ‘Moral Perception’ in Zimmerman, A., Jones, K., and Timmons, M. (eds) *The Routledge Handbook of Moral Epistemology* New York: Routledge pp. 347-360

Ballard, B.S. (2020) ‘The Epistemic Significance of Emotional Experience’ in *Emotion Review*, 13 (2), 113-124

Ben-Ze'ev, A. (2002) ‘Are envy, anger, and resentment moral emotions?’ in *Philosophical Explorations,* 5(2), 148-54

Ben-Ze’ev, A. (2010) ‘The Thing Called Emotion’ in Goldie, P. (ed) *The Oxford Handbook of Philosophy of Emotion* Oxford: OUP pp. 41-62

Bermúdez, J. (1998) *The Paradox of Self-Consciousness* Cambridge, MA: MIT Press

Bradley, F.H. (1886) ‘Is There Any Special Activity of Attention?’ in *Mind* 11 (43), 305-323

Brady, M. (2007) ‘Recalcitrant Emotions and Visual Illusions’ in *American Philosophical Quarterly* 44 (3), 273-284

Brady, M. (2009) ‘The Irrationality of Recalcitrant Emotions’ in *Philosophical Studies* 145 (3), 413-430

Brady, M. (2010) ‘Virtue, Emotion and Attention’ in *Metaphilosophy* 41 (1-2), 115-131

Brady, M. (2013) *Emotional Insight: The Epistemic Role of Emotional Experience* Oxford: OUP

Brewer, B. (1999) *Perception and Reason* Oxford: OUP

Broadbent, D. (1982) ‘Task-Combination and Selective Intake of Information’ in *Acta Psychologica* 50 (3), 253-290

Blackburn, S. (1993) *Essays in Quasi-Realism* Oxford: OUP

Boyd, R. (1988) ‘How to Be a Moral Realist’ in G. Sayre-McCord (ed) *Essays on Moral Realism* Ithaca: Cornell University Press

Blair R. J. R., Mitchell, D., and Blair, K. (2005) *The Psychopath. Emotion and the Brain* Oxford: Blackwell

Block, N. (2007) ‘Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience’ in *Behavioral and Brain Science* 30 (5), 481-548

Brentano, F. (1874) *Psychologie vom empirischen Standpunkt* Hamburg: Felix Meiner. English translation: L. L. McAlister (1973) *Psychology from an Empirical Standpoint* London: Routledge

Brogaard, B. (2008) 'Knowledge-How: A Unified Account' in J. Bengson, M. Moffett (eds.) *Knowing How: Essays on Knowledge, Mind, and Action* Oxford: OUP

Brogaard, B. (2009) 'In Defence of a Perspectival Semantics for 'Know'' in *Australasian Journal of Philosophy* 86 (3), 439-459

Brun, G. and Kuenzle, D. (2008) 'A New Role for Emotions in Epistemology' in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company 1-31

Campbell, J. (2002) *Reference and Consciousness* Oxford: OUP

Calhoun, C. (1984) 'Cognitive Emotions' in C. Calhoun, R. Solomon (eds.) *What is an Emotion: Classic Readings in Philosophical Psychology* Oxford: OUP

Callahan, L. (2018) 'Moral Testimony: A Re-Conceived Understanding Explanation' in *The Philosophical Quarterly* 68 (272), 437-459

Carter, J. and Pritchard, D. (2015) 'Knowledge-How and Cognitive Achievement' *Philosophy and Phenomenological Research* 91 (2015), 181-99

Chappell, S. (2022) *Epiphanies: An Ethics of Experience* Oxford: OUP

Cohon, R. (2018) *Hume's Morality: Feeling and Fabrication* Oxford: OUP

Cowan, R. (2016) 'Epistemic Perceptualism and Neo-Sentimentalist Objections' in *Canadian Journal of Philosophy* 46 (1), 59-81

Cowan, R. (2018) 'Epistemic Sentimentalism and Epistemic Reason-Responsiveness' in A. Bergqvist, R. Cowan (eds.) *Evaluative Perception* Oxford: OUP

D'Arms, J. and Jacobson, D. (2000) 'The Moralistic Fallacy: On the 'Appropriateness' of Emotions' in *Philosophical and Phenomenological Research* 61 (1), 65-90

D'Arms, J. and Jacobson, D. (2003) 'VIII The Significance of Recalcitrant Emotion' in *Royal Institute of Philosophy Supplements* 52: 127-145

D'Arms, J. and Jacobson, D. (2010) 'Demystifying Sensibilities: Sentimental Values and the Instability of Affect' in P. Goldie (ed.) *The Oxford Handbook of Philosophy of Emotion* Oxford: OUP, 585-613

Damasio, A. (1994) *Descartes' Error: Emotion, Rationality and the Human Brain* New York: Putnam

Davidson, D. (1970) 'How is Weakness of the Will Possible' in J. Feinberg (ed.) *Moral Concepts* Oxford: OUP

Deonna, J. and Teroni, F. (2012) *The Emotions: A Philosophical Introduction* Oxford: Routledge

Deonna, J. and Teroni, F. (2015) 'Emotions as Attitudes' in *Dialectica* 69 (3), 293-311

Derryberry, D. and Tucker, D. (1994) ‘Motivating the Focus of Attention’ in P. M. Niedenthal, S. Kitayama (eds) *The Heart’s Eye: Emotional Influence in Perception and Attention* San Diego, CA: Academic Press 167-196

Descartes, R. (1641) *Meditations on First Philosophy* Translated by I. Johnston in A. Bailey (ed.) Broadview Press

Descartes, R. (1649) *The Passions of the Soul* Translated by B. F. Gentile and B. O. Miller in *Foundations of Psychological Thoughts: A History of Psychology* (pp. 5-21) Sage Publications, Inc.

De Sousa, R. (1987) *The Rationality of Emotion* Cambridge, MA: MIT Press

De Sousa, R. (2008) ‘Epistemic Feelings’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company, 185-204

De Sousa, R. and Scarantino, A. (2021) ‘Emotion’ in E. Zalta (eds.) *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition) URL = <https://plato.stanford.edu/archives/sum2021/entries/emotion/> Date accessed 25.2.2023

Döring, S. (2003) ‘Explaining Action by Emotion’ in *The Philosophical Quarterly* 53 (2011), 214-230

Döring, S. (2007) ‘Seeing What To Do: Affective Perception and Rational Motivation’ in *Dialectica* 61 (3), 363-394

Döring, S. (2008) ‘Conflict Without Contradiction’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company 83-104

Dretske, F. (1989) ‘The Need to Know’ in Clay, M. and Lehrer, K. (eds), *Knowledge and Skepticism*, Boulder: Westview Press pp. 89–100

Duncan, J. (1999) ‘Attention’ in R. Wilson, F. Keil (eds.) *The MIT Encyclopedia of the Cognitive Science* Cambridge, MA: MIT Press 39-41

Ekman, P. (1999) ‘Basic Emotions’ in T. Dalgleish, M. Powers (eds.) *Handbook of Cognition and Emotion* Oxford: Blackwell

Elgin, C. (1996) *Considered Judgment* Princeton: Princeton University Press

Elgin, C. (2008) ‘Emotion and Understanding’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company 33-49

Elster, J. (1999) *Alchemies of the Mind: Rationality and the Emotions* Cambridge: CUP

Enoch, D. (2014) ‘A Defense of Moral Deference’ in *Journal of Philosophy* 111 (5) 229-258

Faucher, L. and Tappolet, C. (2008) ‘Introduction: Modularity and the Nature of Emotions’ in L. Faucher and C. Tappolet (eds.), *The Modularity of Emotions*, (*Canadian Journal of Philosophy*, Supplementary Volume 32), vii-xxxi

Fessler, D. (2001) ‘Emotions and the Cost-Benefit Assessment: the role of shame in risk taking’ in G. Gigerenzer and R. Selten (eds) *Bounded Rationality: The Adaptive Toolbox* Cambridge, MA: MIT Press 191-214

Fisher, A. (2014) *Metaethics: An Introduction* Oxford: Routledge

Fletcher, G. (2016) ‘Moral Testimony: Once More with Feeling’ in *Oxford Studies in Metaethics* 11 45-73

Frederickson, B. (1998) ‘Open Hearts Build Lives: Positive Emotions, Induced Through Loving-Kindness Meditation, Build Consequential Personal Resources’ in *Journal of Personality and Social Psychology, 95*(5), 1045–1062

Fredrickson, B. and Branigan, C. (2005) ‘Positive emotions broaden the scope of attention and thought-action repertoires’ in *Cognition and Emotion*, 19(3), 313–332

Frijda, N. (2010) ‘Impulsive Action and Motivation’, *Biological Psychology*, 84(3): 570–579

Gennaro, R. (2004) *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.

Gibbard, A. (1990) *Wise Choices, Apt Feelings: A Theory of Normative Judgment* Harvard University Press

Gigerenzer, G. and Selten, R. (2001) *Bounded Rationality* Cambridge, MA: MIT Press

Goldie, P. (2002) ‘Emotions, Feelings and Intentionality’, *Phenomenology and the Cognitive Sciences*, 1(3): 235–254.

Goldie, P. (2004) ‘Emotion, Reason and Virtue’, in D. Evans and P. Cruse (eds.) *Emotion, Evolution, and Rationality* Oxford: OUP 249–268.

Goldie, P. (2008) ‘Misleading Emotions’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company 149-166

Goldie, P. (2009) ‘Getting Feelings into Emotional Experience in the Right Way’ in *Emotion Review* 1 (3) 232-239

Goldman, A. (1967) ‘A Causal Theory of Knowing’ in *Journal of Philosophy* 64 (12) 357-372

Goldman, A. (1976) ‘Discrimination and Perceptual Knowledge”, The Journal of Philosophy, 73(20): 771–791

Goldman, A. (1986) *Epistemology and Cognition*, Cambridge: Harvard University Press

Gordon 1987

Grzankowski, A. (2020) ‘Navigating Recalcitrant Emotions’ in *Journal of Philosophy* 117 (9) 501-519

Haidt, J. (2003) *The Moral Emotions* in R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (eds.), *Handbook of affective sciences* (pp. 852–870). Oxford: OUP.

Hills, A. (2009) ‘Moral Testimony and Moral Epistemology’ in *Ethics* 120, 94–127.

Hills, A. (2016) ‘Understanding Why’ in *Nous* 50 (4) 661-688

Hookway C. (2003) ‘Affective States and Epistemic Immediacy’ in *Metaphilosophy*, 34 (1-2), 78-96

Hookway C. (2008) ‘Epistemic Immediacy, Doubt and Anxiety: On a Role for Affective States in Epistemic Evaluation’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) *Epistemology and Emotions* Aldershot: Ashgate Publishing Company 51-66

Hopkins, R. (2007) ‘What is Wrong with Moral Testimony’ in *Philosophy and Phenomenological Research* 74 (3):611-634

Howell, R. (2014) ‘Google Morals, Virtue, and the Asymmetry of Deference’ in *Noûs* 48 (3) 389-415

Hume ([1739/40], 276, 1777 Hume treatise of Human nature

Hursthouse, R. (1991) ‘Arational Actions’ in *Journal of Philosophy* 88 (2) 57-68

Ichikawa, J. and Steup, M. ‘The Analysis of Knowledge’ in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition) URL = <https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/> Date accessed 19.1.2023

Jennings, C. (2012) ‘The Subject of Attention’ in *Synthese* 189 (3):535-554

Jenkins, C. (2015) ‘What Quasi-Realists Can Say About Knowledge’ in R. Johnson and M. Smith (eds.) *Passions and Projections: Themes from the Philosophy of Simon Blackburn* Oxford: OUP 64-84

Johnson, A. and Proctor, R. (2004) *Attention: Theory and Practice* Sage Publications

Johnston M. (2001) ‘The Authority of Affect’ in *Philosophy and Phenomenological Research* 63 (1):181-214

Joyce, R. (2000) *The Evolution of Morality* Cambridge, MA: MIT Press

Kenny, A. (1963) *Action, Emotion and Will* New York: Humanities Press

Kriegel, U. (2009) *Subjective Consciousness* Oxford: OUP

Kriegel, U. (2019) ‘The Intentional Structure of Moods’ in Philosophers’ Imprint 19 (49) 1-19

Kvanvig, J. (2003) *The Value of Knowledge and the Pursuit of Understanding* Cambridge: CUP

Lamme, V. (2003) ‘Why Visual Attention and Awareness are Different’ in *Trends in Cognitive Sciences* 7 (1) 12-18

Lance, M. and Tanesini, A. (2004) ‘Emotion and Rationality’ in *Canadian Journal of Philosophy* 34 (Supplement) 275-295

Lazarus, R. (1991) *Emotion and Adaptation* Oxford: OUP

LeDoux, J. (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* New York: Simon and Schuster.

Lewis, D. (1988) ‘What Experience Teaches’ in W. Lycan (ed.) *Mind and Cognition* Oxford: Blackwell 29-57

Lord, E. (2016) ‘On the Rational Power of Aesthetic Testimony’ in *British Journal of Aesthetics* 56 (1) 1-13

Lutz, M. and Lenman, J. (2021) ‘Moral Naturalism’ in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), URL =

<https://plato.stanford.edu/archives/win2021/entries/naturalism-moral/> Date accessed 30.3.2023

Mack, A. and Rock, I. (1998) *Inattentional Blindness* Cambridge, MA: MIT Press

Mack, A. and Rock, I. (2003) 'Inattentional blindness: An overview' in Current Directions in Psychological Science 12 (5):180-184

Mackie, J. L. (1977) *Ethics: Inventing Right and Wrong* Penguin Books

Maibom, H. (2020) *Empathy* Oxford: Routledge

McDowell, J. (1979) 'Virtue and Reason' in *The Monist* 62 (3) 331-350

McDowell, J. (1985) 'Values and Secondary Qualities' in T. Honderich (ed.), *Morality and Objectivity* London: Routledge 110-129

McDowell, J. (1994) 'The Content of Perceptual Experience' in Philosopical Quarterly 44 (175) 190-205

Mason, M. (2003). 'Contempt as a Moral Attitude' *Ethics* 113 (2): 234-72.

McGrath, S. (2009) 'The Puzzle of Pure Moral Deference' *Philosophical Perspectives*, 23 321-344

McGrath, S. (2011) 'Skepticism about Moral Expertise as a Puzzle for Moral Realism' *The Journal of Philosophy*, 108 (2011): 111-137.

McGrath, S. (2019) *Moral Knowledge* Oxford: OUP

Mitchell J. "The Epistemology of Emotional Experience," *Dialectica*, Vol. 71, N 1 (2017): 57-84

Mole, C. (2010) *Attention Is Cognitive Unison: An Essay in Philosophical Psychology* Oxford: OUP

Mole, C. (2021) 'Attention' in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), URL = <https://plato.stanford.edu/archives/win2021/entries/attention/> Date accessed: 12.10.2022

Moore, C. and Dunham, P. (eds.) (1995) *Joint Attention: Its Origins and Role in Development* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Most S. B., Chun M. M., Widders D. M., Zald D. H. (2005). 'Attentional rubbernecking: cognitive control and personality in emotion-induced blindness' Psychonomic Bulletin and Review 12 654–661

Neisser, U. (1976) *Cognition and Reality,* San Francisco: Freeman

Nickel, P. (2001) 'Moral testimony and its authority' in *Ethical Theory and Moral Practice* 4 (3) 253-266

Nisbett, R. and Ross, L. (1980) *Human Inference: Strategies and Shortcomings of Social Judgment* Englewood Cliffs, NJ, USA: Prentice-Hall

Nozick, R. (1981) *Philosophical Explanations*, Cambridge, MA: Harvard University Press.

Nussbaum, M. (2001) *Upheavals of Thought: The Intelligence of Emotions* Cambridge: CUP

Oddie, G. (2005) *Value, Reality, and Desire* Oxford: Clarendon Press

Olson, J. (2014) *Moral Error Theory: History, Critique, Defence* Oxford: OUP

Pashler, H. (1998) *The Psychology of Attention* Cambridge, MA: MIT Press

Peacocke, H. (1989) 'Perceptual content' in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan* New York: OUP.

Peters, E. (2006) 'Affect and decision making: a "hot" topic' *in Journal of Behavioral Decision Making* 19 (2) 79-85

Pillsbury, W. (1973) *Attention* Arno Press

Platts, M. (1979) 'Ways of Meaning' in *Linguistics and Philosophy* 4 (1):141-156

Price, C. (2015) *Emotion* Malden, Massachusetts: Polity Press

Prinz, J. (2004) *Gut Reactions: A Perceptual Theory of Emotion* New York: Oxford University Press

Prinz, J. (2007) *The Emotional Construction of Morals* Oxford: Oxford University Press

Prinz, J. and Nichols, S. (2010) 'Moral Emotions' in J. Doris and the Moral Psychology Research Group (Eds.), *The Moral Psychology Handbook*, Oxford: OUP

Pritchard, D. (2008) 'Radical Scepticism, Epistemic Luck and Epistemic Value', *Proceedings and Addresses of the Aristotelian Society* (suppl. vol.) 82 (2008), 19-41

Railton, P. (2014) 'The Affective Dog and Its Rational Tale: Intuition and Attunement' in *Ethics* 124 (4) 813-859

Raleigh, T. (2019) 'Introduction: The Recent Renaissance of Acquaintance' in J. Knowles, T. Raleigh (eds.) *Acquaintance: New Essays* in Oxford: OUP

Riaz, A. (2015) 'Moral understanding and knowledge' in Philosophical Studies 172 (1):113-128

Roberts, D. (2013) 'Thick Concepts' in *Philosophy Compass* 8 (8) 677-688

Roush, S. (2017) 'Epistemic Self-Doubt' in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), URL = <https://plato.stanford.edu/archives/win2017/entries/epistemic-self-doubt/> Date accessed: 4.2.2023

Ryle, G. (1949) *The Concept of Mind* London: Hutchinson

Salmela, M. (2011) 'Can emotion be modelled on perception?' in *Dialectica* 65 (1) 1-29

Scanlon, T. (1998) *What We Owe to Each Other* Belknap Press of Harvard University Press

Scarantino, A. (2015) 'Basic Emotions, Psychological Construction and the Problem of Variability': 334–376

Scherer, K. (2001) ‘Appraisal Considered as a Process of Multilevel Sequential Checking’, in K. Scherer, A. Schorr, and T. Johnstone (eds.) *Appraisal processes in emotion: Theory, methods, research* Oxford: OUP 92–120

Schroeder, M. (2007) *Slaves of the Passions* Oxford: OUP

Schroeder, T. (2020) ‘Desire’ in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), URL = <https://plato.stanford.edu/archives/sum2020/entries/desire/> Date accessed 30.3.2023

Shafer-Landau, R. (2003) *Moral Realism: A Defence* Oxford: OUP

Sinclair, N. (2012) ‘Expressivist Explanations’ in Journal of Moral Philosophy 9 (2):147-177

Sinclair, N. (2021) *Practical Expressivism* Oxford: OUP

Sinclair, N. and Chamberlain, J. (2022) ‘The Evolutionary Debunking of Quasi-Realism’ in Diego E. Machuca (ed.), *Evolutionary Debunking Arguments: Ethics, Philosophy of Religion, Philosophy of Mathematics, Metaphysics, and Epistemology* London: Routledge 33-55

Sliwa, P. (2015) ‘IV – Understanding and Knowing’ in Proceedings of the Aristotelian Society 115 (1pt1) 57-74

Sliwa, P. (2017) ‘Moral Understanding as Knowing Right from Wrong’ in *Ethics* 127 (3) 521-552

Slovic, P. (2007) ‘Psychic Numbing and Genocide’ in *APA Psychological Science Agenda*, 21 (10) URL: < https://www.apa.org/science/about/psa/2007/11/slovic> Date accessed: 20.3.2023

Smallwood, J. and Schooler, J. (2009) ‘Mind-Wandering.’ T. Bayne, A. Cleermans and P. Wilken (eds.) *The Oxford Companion to Consciousness* Oxford⁄New York: OUP 143–5.

Smith, M., Lewis, D., and Johnston, M. ‘Dispositional Theories of Value’ in *Aristotelian Society Supplementary* Volume 63 (1):89-174

Smithies, D. (2011)‘Attention is Rational-Access Consciousness.’ *Attention: Philosophical and Psychological Essays*. C. Mole, D. Smithies and W. Wu (eds.). Oxford⁄New York: OUP 247–73.

Solomon, R. (1973) ‘Emotions and Choice’ in *Review of Metaphysics* 27 (1):20 - 41

Srneevisan, G. (2020) *Emotion and Virtue* Princeton University Press

Stampe, D. (1987) ‘The Authority of Desire’ in Philosophical Review 96 (July):335-81

Stanley, J. and Williamson, T. (2001) ‘Knowing How’ in Journal of Philosophy 98 (8):411-444

Starkey, C. (2008) ‘Emotion and Full Understanding’ in Ethical Theory and Moral Practice 11 (4):425-454

Steup, M. (2001) ‘The Analysis of Knowledge’ in E. Zalta (ed.) The Stanford Encyclopedia of Philosophy (Spring 2001 Edition) URL = <

https://plato.stanford.edu/archives/spr2001/entries/knowledge-analysis/> Date accessed 19.1.2023

Stepper, S. and Strack, F. (1993) 'Proprioceptive Determinants of Emotional and Nonemotional Feelings' in *Journal of Personality and Social Psychology* 64(2):211-220

Stocker, M. (1983) 'Psychic feelings: Their importance and irreducibility' in *Australasian Journal of Philosophy* 61 (1):5-26

Stratton-Lake, P. (2020) 'Intuitionism in Ethics' in E. Zalta (eds.) *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), URL = <https://plato.stanford.edu/archives/sum2020/entries/intuitionism-ethics/> Date accessed 30.3.2023

Streumer and Wodak 2021 Streumer B. and Wodak D., "Why Formal Objections to the Error Theory Fail" in *Analysis* Vol. 81, (April 2021): 254-262.

Strevens, M. (2013) 'No understanding without explanation' in Studies in History and Philosophy of Science Part A 44 (3):510-515

Stroud, B. (2000) *Understanding Human Knowledge: Philosophical Essays* Oxford: OUP

Stueber, K. (2019) 'Davidson, Reasons, and Causes: A Plea for a Little Bit More Empathy' in *Journal for the History of Analytical Philosophy* 7 (2):59-75

Swain, M. (1981) *Reasons and Knowledge* Ithaca: Cornell University Press

Tanesini, A. (2008) 'Self-Knowledge and Resentment' in *Philosophical Books* 49 (3):238-245

Tappolet, C. (2012) 'Emotions, perceptions, and emotional illusions' in C. Clotilde (ed.), Perceptual Illusions. Philosophical and Psychological Essays, Palgrave-Macmillan. pp. 207-24

Tappolet, C. (2016) *Emotions, Value, and Agency* Oxford: UP

Thagard, P. (2002) '*The passionate scientist: Emotion in scientific cognition*' in Peter Carruthers, Stephen P. Stich & Michael Siegal (eds.), The Cognitive Basis of Science. Cambridge: Cambridge University Press. pp. 235-261

Thagard, P. (2008) *Hot thought: Mechanisms and applications of emotional cognition* Cambridge, MA: MIT Press

Tooby, J., and Cosmides, L. (1990) 'The past explains the present: Emotional adaptations and the structure of ancestral environments' *Ethology & Sociobiology*, 11(4-5), 375–424

Treisman, A. (1964) 'Selective Attention in Man' in British Medical Bulletin 20 (1):12-16

Treisman, A. (2006) 'How the deployment of attention determines what we see' *Visual Cognition* 14: 411-443.

Vanello, D. (2020) 'Affective Disclosure of Value: emotional experience, neo-sentimentalism and learning to value' in Philosophy 95 (3): 261-283

Van Roojen, M. (2004) ‘Moral Cognitivism vs. Non-Cognitivism’ in E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), URL = <https://plato.stanford.edu/archives/fall2018/entries/moral-cognitivism/> Date accessed: 12.12.2022

Watzl, S. (2010) *The Significance of Attention* PhD Thesis, Columbia University, 2010.

Waltz, S. (2011) ‘The Nature of Attention’ in *Philosophy Compass* 6/11 (2011): 842–853

Whiting, D. (2015) ‘The Glass is Half Empty: A New Argument for Pessimism about Aesthetic Testimony’ in British Journal of Aesthetics 55 (1):91-107

Wild, M. (2008) ‘Critical Reflections on Affective Epistemology’ in G. Brun, U. Dogluoglu, D. Kuenzle (eds.) Epistemology and Emotions Aldershot: Ashgate Publishing Company 125-149

Williams, M. (2001) *Problems of Knowledge* Oxford: OUP

Wodak, D. (2019) ‘Approving on the Basis of Moral and Aesthetic Testimony’ in *Oxford Studies in Metaethics* vol. 16 Oxford: OUP 183-206

Wu, W. (2011) ‘What is Conscious Attention?’ in *Philosophy and Phenomenological Research* 82 (1):93-120

Zagzebski, L. (2001) ‘Recovering Understanding’ in M. Steup (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: OUP