

RETHINKING THE FITTINGNESS OF EMOTIONS

A Dissertation

by

DONG AN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Linda Radzik
Co-Chair of Committee,	Martin Peterson
Committee Members,	Kenny Easwaran
	José Luis Bermúdez
	Heather Lench
Head of Department,	Theodore George

May 2021

Major Subject: Philosophy

Copyright 2021 Dong An

ABSTRACT

Emotions are always *about* something and are thus intentional. Emotional fittingness is a normative concept that is used to describe the relationship between an emotion and its object. If an emotion is fitting, it must somehow correctly capture the connection between the particular object and the formal object, i.e., the corresponding evaluative properties.

It is widely held in the literature that there is a generic standard of correctness with respect to the evaluative property that characterizes fittingness. I accept that there is such a notion of fittingness that is distinct. However, I argue that three constraints need to be satisfied in an account of emotional fittingness.

First, a notion of fittingness should leave room for metaethical disagreements and thus should not assume that the criterion of fittingness is universally applicable to everybody in the same sort of situation. Second, although fittingness assessment is orthogonal to moral justification, as indicated by the argument against the “moralistic fallacy,” the judgment of whether an emotion is fitting sometimes is constrained by background moral considerations. Third, one important value that has been attributed to an emotion’s being fitting is actually constituted by something beyond fittingness, namely by (affective) appreciation, which is fitting emotion plus understanding. Thus, an account of fittingness should not confuse experiencing a fitting emotion towards the object with appreciating the object.

Fittingness is a normative dimension that is unique yet limited and constrained. While I do not develop a full account of emotional fittingness, the constraints argued for here should apply to all accounts of fittingness.

ACKNOWLEDGEMENTS

I am very grateful for the comments, advice, and support from many people throughout this long time of dissertation writing.

I am very lucky to have a generous and caring committee. First and foremost, I would like to express my deep gratitude to Linda Radzik. She kindly agreed to let me audit her class even before I was officially enrolled in the Ph.D. program. Her suggestions and encouragement not only helped me transition to the American educational setting, which was all new to a student coming from the other side of the Earth, but accompanied me throughout my studies. When I struggled and felt desperate, her approach to and attitude towards philosophy reminded me of what made me interested in philosophy at the very beginning. She is not only an advisor of my dissertation, but also a role model of how to live a philosophical life in the meantime not to be overwhelmed by philosophy.

I would also deeply thank my co-advisor, Martin Peterson. I could not have carried on without his constantly insightful feedback on many drafts that made or did not make to the final dissertation. His standards on clarity, precision, and using specific examples to illustrate relevant points all help me tremendously with sharpening my philosophical skills.

I owe a great deal to my committee members. Thanks to Kenny Easwaran for his valuable suggestions and reading recommendations. In many cases, his comments helpfully shifted the direction of my project. Thanks to José Luis Bermúdez for his

advice and allowing me to participate in his two projects, which offered tremendous resources related to my dissertation topic and greatly broadened my horizon. Thanks to Heather Lench for her work on affective science. I was first introduced to the psychological research of emotions by her and learned a lot from her. I could not have a better committee.

Other faculty members at the Texas A&M University have helped me over the years. In particular, thanks to Ted George for conversations and reading advice; to Nathan Howard for discussions and comments on my draft; to Clare Palmer for offering me opportunities and suggestions on my professional career; to Kristi Sweet, Gary Varner, and Emily Brady for their valuable comments on a talk I gave on one chapter of the dissertation; and to Glen Miller for helping me with teaching advice.

I would like to thank my fellow Ph.D. students in the department. Thanks to Wendy Bustamante, who was my student mentor (the best one) when I was a first-year student, for helping me navigating the program, answering my questions whenever I had, and listening to my complaints. Thanks to Jennifer Wargin, who was willing to form a useful virtual dissertation group when both of us had to finish remotely. Thanks to Rob Reed, my last year's office mate, who shared his view on my dissertation topic and exchanged ideas for many times. Thanks to Philip Cho, Steve Dezort, and Ryan Manley, for participating in a survey study related to my project.

I also would like to thank the generous help from people I met during conferences, workshops, and fellowship opportunities. Special thanks to the Glasscock Center at Texas A&M University for the helpful one-year fellowship; to Preston Werner,

Christine Tappolet, and Mark Thomson in the Emotion and Perception in Moral Epistemology summer fellowship at Hebrew University at Jerusalem; to Zhen Cai and her Ethic workshop at East China Normal University; to Xiangdong Xu and his workshop on Knowledge, Value, and Virtues at Zhejiang University; to Fabrice Teroni, Julien Deonna, Ronald de Sousa, Trip Glazer, Jonas Blatter, and Kris Goffin.

Special thanks to Kelly Clark, who led the “Virtues and Values” program. It is through that program that I first got the opportunity to study for a year at Baylor University and got to see how doing philosophy is like in the English-speaking world. That experience shaped my thoughts about getting a career in philosophy and prepared me to get into the Ph.D. program at Texas A&M.

Many thanks to my parents for their unconditional support and understanding of me not being able to visit them often, especially my mother, who was willing to act as my student to give me feedback when I did not have any real student to teach philosophy. Thanks to my daughter, Tracy, for bringing incalculable joy to me and unexpected inspiration to both my philosophical work and real life. Finally, a profound thank you to Xing Wang, my precious partner and co-parent, who traveled with me from our undergraduate time in China to our graduate time in United States, who endured my emotional turmoil during many difficult times, who supported me, trusted me, and was proud of me when I was in deep crisis of self-worth. My family was the reason why I chose to work on philosophy of emotion. Now that the dissertation journey has come to an end, I think I understand better emotions as well as our relationships.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professors Linda Radzik (chair), Martin Peterson (co-chair), Kenny Easwaran, and José Luis Bermúdez of the Department of Philosophy and Professor Heather Lench of the Department of Psychological and Brain Sciences.

The section 4 in Chapter 3 were partly contributed by Kaiyuan Chen of the Department of Psychology at Claremont Graduate School and were published in 2021.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGEMENTS.....	iv
CONTRIBUTORS AND FUNDING SOURCES	vii
TABLE OF CONTENTS	viii
CHAPTER I INTRODUCTION.....	1
1.1. Introducing Fittingness	1
1.2. The Significance of Fittingness.....	3
1.2.1. Thought Experiments	5
1.2.2. The Value of Emotions.....	8
1.2.3. Emotion and Reason	16
1.2.4. Explaining Value.....	19
1.3. What Is Fittingness?.....	19
1.3.1. Reasons-based Accounts	20
1.3.2. Value-based Accounts.....	23
1.3.3. Correct Representation	24
1.3.4 Fittingness First.....	25
1.4. Summary of the Chapters.....	26
1.5. Conclusion.....	29
CHAPTER II AGAINST THE UNIVERSALIST ASSUMPTION OF EMOTIONAL FITTINGNESS	30
2.1. Introduction	30
2.2. The Standard Model.....	32
2.3. The Universalist Assumption.....	36
2.3.1. D’Arms and Jacobson’s Rational Neo-Sentimentalism	36
2.3.2. Tappolet’s Representational Neo-Sentimentalism.....	40
2.4. Cases of Varied Intuitions.....	44
2.5. What to Make of the Disagreement?	51
2.6. Conclusion.....	56
CHAPTER III JOKES CAN FAIL TO BE FUNNY BECAUSE THEY ARE IMMORAL: THE INCOMPATIBILITY OF EMOTIONS.....	58

3.1. Introduction.....	58
3.2. The “Moralistic Fallacy” and a Response.....	60
3.3. Avoiding the Fallacy.....	64
3.4. Moral Reason: Intramural vs. Extramural.....	65
3.5. Negative Emotions with High Arousal: Moral Disgust and Anger.....	70
3.6. Incompatibility of Emotions.....	72
3.6.1. The Theoretical Point of View.....	73
3.6.2. The Empirical Point of View.....	76
3.7. “Can” and “Is”.....	80
3.8. Putting Things Together.....	81
3.9. Conclusion.....	84
 CHAPTER IV APPRECIATION AS AN EPISTEMIC EMOTION: HOW FITTING ANGER IS DIFFERENT FROM INJUSTICE APPRECIATION	 88
4.1. Introduction.....	88
4.2. Understanding as Grasping.....	91
4.3. Appreciation as an Epistemic Emotion.....	97
4.4. Assessing Fitting Emotions.....	103
4.4.1. A linguistic clarification.....	105
4.4.2. Fitting anger without appreciation.....	107
4.5. The Many Kinds of Appreciation.....	114
4.5.1. Aesthetic, moral, and epistemic appreciation.....	114
4.5.2. Appreciation requires emotion as a prerequisite.....	116
4.6. Aesthetic Appreciation and Aesthetic Emotions.....	121
4.7. Conclusion.....	124
 CHAPTER V CONCLUSIONS	 126
 REFERENCES	 135

CHAPTER I

INTRODUCTION

1.1. Introducing Fittingness

My 4-year-old daughter is extremely shy. She avoids talking to people whenever she can. One day, when she was at preschool, she pooped in her pants. But she was too shy and embarrassed to tell the teachers. Unfortunately, because of the COVID-19 pandemic, all teachers wore masks, so no one noticed the smell. She ended up wearing the dirty pants and played for two hours as if nothing happened until I picked her up at the end of the day. I was both amazed and amused. Of course, I did not express these emotions, for that would have embarrassed her even more.

The rest of the story is nothing more than the gentle preaching a mother often gives to her child. However, the point of this anecdote is that there are quite a few emotions that we can say are fitting or not fitting. We can question whether it is fitting for my daughter to feel timid to talk to people, whether it is fitting for her to feel embarrassed to tell her teachers about the accident, and whether it is fitting for me to feel amazed and amused by all this. We feel emotions often. We evaluate the fittingness of our emotions almost as often. This dissertation investigates the notion of emotional fittingness.

As a technical term, emotional fittingness is a normative concept that is used to describe the relationship between an emotion and its object. Emotions are always *about* something. My daughter's timidity is about talking to people. Her embarrassment is

about her having an accident. My amazement is about her being so tolerant of the accident. My amusement is about her funny way of handling the accident. These specific things are called the “particular objects” of emotions. In feeling timid, my daughter found talking to people intimidating. In feeling embarrassment, she found her having an accident embarrassing. In feeling amazed, I found her being tolerant of the accident amazing. In feeling amused, I found her way of handling the accident amusing. The words “intimidating,” “embarrassing,” “amazing,” and “amusing” all refer to certain kinds of evaluative properties. These evaluative properties are called the “formal objects” of the corresponding emotions. If an emotion is fitting, it must somehow correctly capture the connection between the particular object and the formal object.

In my dissertation, I propose three constraints that an account of fittingness should satisfy. It seems that everyone in the literature agrees that there is a generic standard of correctness with respect to the evaluative property that characterizes fittingness. I accept that there is such a notion of fittingness. However, I argue that three constraints need to be satisfied in an account of emotional fittingness.

First, a notion of fittingness should leave room for metaethical disagreements and thus should not assume that criteria for the fittingness of particular emotions are universally applicable to everyone. Second, although fittingness assessment is orthogonal to moral justification, as indicated by the argument against the “moralistic fallacy,” the judgment of whether an emotion is fitting will sometimes be constrained by background moral considerations. Third, one important value that has been attributed to an emotion’s being fitting is actually constituted by something beyond its fittingness,

namely by (affective) appreciation, which is fitting emotion plus understanding. Thus, an account of fittingness should not confuse experiencing a fitting emotion towards an object with appreciating the object.

Fittingness is a normative dimension that is unique yet limited and constrained. Although I do not intend to develop a full account of emotional fittingness, the constraints argued for here should apply to all accounts of fittingness.

1.2. The Significance of Fittingness

Surely, the fittingness assessment is not the only way we can evaluate our emotions. As my opening example shows, it is prudentially inappropriate for me to show or feel the amazement and amusement in front of my daughter, because that would make her feel more embarrassed. These emotions can even be morally inappropriate, if they hurt her feelings. But this prudential or moral evaluation is different from the fittingness evaluation. Fittingness is proposed to capture the relationship between the particular object and the formal object of an emotion. Thus, even though it is prudentially or morally inappropriate to feel those emotions, they nevertheless can be fitting, as long as they match their objects correctly.

Similarly, we can also evaluate emotions in other ways irrelevant to fittingness. For example, if, when my daughter grows up, she still feels embarrassed by the incident when I joke about it, then perhaps we should say that her embarrassment lasts too long to be appropriate. Or, if, as an adult, she feels even more embarrassed by the incident, we probably should say that her embarrassment is too much to be appropriate. No matter

how appropriate or inappropriate these emotions are in terms of duration and intensity, these assessments are separable from the fittingness assessment.

The prudential, moral, duration, and intensity assessments of emotions are common and important in their own ways. But they should not be confused with the fittingness assessment, which is specifically about the emotional objects. Neither should their importance undermine the importance of the fittingness assessment. They register different aspects of our emotional experiences, and thus can be discussed relatively independent from each other.

The importance of the concept of fittingness can be discussed from at least four perspectives. First, many convincing thought experiments show that the fittingness evaluation applies broadly, not only to emotions, but also to beliefs and intentions. Thus, *prima facie*, there is a task to explain this seemingly unified phenomenon. Second, in the past several decades, the positive role of emotions has been more and more recognized, not only in philosophy but also in psychology. Many such discussions take fitting emotions as the subject matter. To understand these arguments, we should first understand the idea of fittingness. Third and related, these discussions offer us precious opportunities to reevaluate the traditional dichotomy between emotion and reason. Thus, understanding fittingness and fitting emotions opens a door to reassess many resources in the history of philosophy. Finally, fittingness has been a useful construct in value theory. It is deployed to explain the concept of value. Given the importance of the notion of value, fittingness naturally inherits this importance. In the following, I will explain these aspects in more detail.

1.2.1. Thought Experiments

In the fittingness literature, three thought experiments stand out as particularly significant. The first is the “immoral funny joke” (D’Arms & Jacobson, 2000a). Suppose that you hear a clever sexist joke. You find the joke quite funny just in terms of being a joke. However, you also think that the joke is horrible because it is based on some extremely sexist bias. The creators of this example, Justin D’Arms and Daniel Jacobson, and almost every other philosopher writing on the topic, all agree that although it is morally inappropriate to feel amused by the joke, it is absolutely fitting to be amused by it just in terms of its funniness. This conclusion seems to be in accordance with our considered intuition. It seems that we do feel that there is a kind of assessment about the funniness that is independent from our moral evaluation of it.¹

The second example is the “enviable tenured colleague” (D’Arms & Jacobson, 2000a). Suppose that you are a junior tenure-track faculty member in your department. Susan is your colleague and she just won tenure. Prudentially speaking, it is not appropriate to envy her, since doing that would potentially jeopardize your relations with her and eventually you will need her vote. However, it is fitting to envy her because her tenure is indeed enviable. Again, it seems intuitive that there is a kind of assessment about whether Susan’s achievement is worthy of envy regardless of whether it is prudentially appropriate or not to feel the emotion.

¹ However, my chapter 3 shows a picture of amusement and morality more complicated than this.

The third example is the “evil demon” (Rabinowicz & Rønnow-Rasmussen, 2004). Suppose that an evil demon asks you to admire him, otherwise he will inflict severe pain on your family. Presumably, you have a reason to admire the evil demon. We may interpret this reason either as a moral one or a prudential one. However, it is not fitting to admire the evil demon, since evil demons do not have the properties required for being worthy of admiration.

The examples of “immoral funny joke” and “enviable tenured colleague” show that moral or prudential considerations are not necessary for fittingness evaluations. It is easy to construct examples involving other emotions with the same structure, such as my anger being fitting yet destructive, her grief being fitting but imprudent, his anxiety being fitting but unhealthy. This indicates that it is a fairly common phenomenon. The “evil demon” example shows that moral or prudential considerations are not sufficient for fittingness evaluations. It is even easier to construct examples involving other emotions with the same structure. We can just change “admiration” to any other emotion: the evil demon asks you to love him, hate him, feel proud of him, feel sad about him, etc. No matter whether we have relevant reasons to have these emotions towards the evil demon, the mere fact that he asks us to do so is not the right kind of reason for these emotions. Thus, whatever we want to use to characterize this concept of fittingness, it seems to be distinct from these other considerations.

This phenomenon is not restricted to emotions but generalizes to other attitudes such as beliefs and intentions (Gertken & Kiesewetter, 2017). On the one hand, we can modify the evil demon example to come up with the corresponding examples in beliefs

and intentions. For example, suppose that the evil demon asks you to believe that this year is 1896, otherwise he will inflict severe pain on your family. Presumably, you have a pragmatic reason to believe that proposition. However, it is not fitting to believe it in terms of its credibility or truth. The notion of fittingness in beliefs seems to be related to credibility or truth, which are concepts that are related to the epistemic nature of the attitude of belief. This indicates that the notion of fittingness in emotions may also have to do with the nature of the particular emotions in question. Similarly, we can imagine the evil demon, or an eccentric billionaire as in Gregory Kavka's example, offers you a large reward for intending to drink a glass of toxin, regardless of whether you actually drink it or not (Kavka, 1983). It seems that you have a pragmatic reason to intend to drink it. However, it is not fitting to form the intention. Here, we make this judgment because the nature of intention is action-related. Just as in beliefs and emotions, the notion of fittingness in intentions is sensitive to the nature of the attitude. These examples show that even if there are prudential considerations for an attitude, they do not necessarily make the attitude fitting.

On the other hand, we can imagine cases where an attitude is fitting even though there are other considerations against the attitude. For example, Bruno believes that the Earth goes around the Sun. Although there are strong religious and prudential reasons for him not to believe so, his belief is still fitting in terms of its credibility or truth. Similarly, we can imagine that a father whose young daughter has gone missing intends to stop every other activity and devote all his life to looking for her. In this case, he may

lack a strong prudential reason to intend such, since the decision is highly likely to lead to a miserable life. However, it is hard to criticize his intention as not fitting.

These thought experiments clearly show that in general, there is a type of attitude assessment independent of morality, prohibitions, or incentives. We call it the fittingness assessment. We know it is related to the nature of the attitude in question. The task is to explicate this relation.

1.2.2. The Value of Emotions

Emotion research has seen a tremendous growth in the past several decades both in psychology and in philosophy. Scholars have been working on theorizing and testing the positive role of emotions. Psychologists argue that emotions function as signals and motivate us to act accordingly. Philosophers argue that emotions have the instrumental epistemic value of offering information, the non-instrumental epistemic value of constituting evaluative knowledge, the instrumental moral value of helping build characters, and the non-instrumental moral value of constituting moral sensitivity. I discuss each of these in detail below.

1.2.2.1. Psychological Research: Signaling and Motivating

Psychologists have pointed out many different roles emotions can play. Two important roles are relevant to our discussion of fittingness: signaling and motivating. According to functionalist theories, emotions are adapted reactions to specific goal-related situations (Ekman, 1992; Lazarus, 1991; Lench, Bench, Darbor, & Moore, 2015; Lench, Flores, & Bench, 2011; Levenson, 2014; Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005; Pinker, 1997; Scherer, 1984). Scientific studies usually focus on the function rather than

the value of emotions. However, in many cases, it is reasonable to argue that emotions are valuable because of their functions, i.e., they are mechanisms to help achieving our goals.

First, through emotions, we can detect evaluative properties. For example, anger signals that something offensive has occurred, fear signals that the situation is dangerous, boredom signals that the current situation is no longer stimulating. It is argued that although it is possible to detect these values without emotions, emotions can improve the speed at which we notice them and thus serve as shortcuts that require lower cognitive costs. Second, emotions motivate agents to take corresponding actions in response to the evaluative properties. For example, anger can activate the readiness for action to remove obstacles to goals, fear can motivate us to flee or fight back, boredom can encourage us to pursue goals that are more interesting.

Thus, emotions are useful tools for navigating the environment by providing us with information and urging us to act. But emotions can have such an instrumental value only if they reliably lead us in the correct direction. For example, if I feel angry whenever I see the color red, then the anger provides me with the misleading information and so likely motivates me to act inappropriately. In this case, anger does not promote the goals I most likely have and thus is not useful. This is an example of unfitting emotion. Unfitting emotions lose their value in offering accurate information and motivating appropriate action. Thus, at least in understanding the value of emotions in these two senses, we need to be clear about what makes an emotion fitting.

1.2.2.2. Philosophical Research: Epistemic Value

In contemporary philosophy of emotion, theorization has gone through four stages. First, the Jamesian feeling theory says that emotions are just feelings of bodily changes. The slogan coined by William James is, “We feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful” (James, 1884, 190).

This theory has been largely abandoned due to the consensus among philosophers that it cannot explain the intentionality of emotions. That is, as we mentioned above, emotions are *about* something. Although bodily changes saliently accompany emotions, they do not tell us why certain situations elicit only certain emotions. Accordingly, the feeling theory cannot distinguish different types of emotions. However, we do think that, for example, anger is different from sadness, although they may accompany similar bodily responses such as crying. Thus, the theory proposed to replace the feeling theory focuses on the intentional and thus the cognitive aspect of emotions. Cognitivists argue that in feeling a certain emotion we form a kind of judgment (Solomon, 1984; Nussbaum, 2001). This judgment is our appraisal of the situation. For example, in feeling sad, we judge the situation as instantiating an irrevocable loss.

Being able to explain the intentionality of emotions is the chief advantage of the cognitivist theory. However, the theory has been criticized as being over intellectualized since explicit judgment is not necessary for forming an emotion (Deigh, 1994; Goldie, 2000; Greenspan, 1988; Pugmire, 1998). For example, infants who do not have the

cognitive capacity to form a judgment can nevertheless feel sad, angry, or happy. Thus, the perceptualist theory has been proposed as a revision and it has become the mainstream today (de Sousa, 1987a; Goldie, 2000; Helm, 2001; Prinz, 2004; Roberts, 2003; Tappolet, 2016). According to this theory, emotions are intentional as the cognitivists suggest. But rather than being a kind of judgment, which is explicitly cognitive and conceptual, emotions are more like perceptions, which can have non-conceptual intentional content. In this picture, emotions are still evaluations. But they represent the evaluative properties just like perceptions represent the perceptual properties.

Although the perceptualist theory has advantages over the feeling and the cognitive theories, it still has difficulties. Philosophers who are unconvinced by the analogy between emotion and perception propose the attitudinal theory, suggesting that emotions are *sui generis* and thus not reducible to other attitudes (Deonna & Teroni, 2012, 2015a). In this theory, emotions remain evaluative. Here, different types of emotions are different not because they have different representational contents, but because they are different attitudes. Emotions are evaluative attitudes.

As this quick review of the theories of emotion shows, it is the consensus among philosophers that emotions have evaluative properties as their intentional objects. Although it is still highly contentious how best to explain this emotional intentionality, almost everyone agrees that because of this intentionality, emotions have important epistemic value. In other words, in having an emotion, we get to know the corresponding evaluative property. For example, in feeling fear, sadness, anger, and admiration, we

evaluate the object as fearsome, a loss, unjust, and admirable, respectively. In this sense, emotions are our epistemic tools for detecting evaluative properties, just as perceptions are our epistemic tools for detecting perceptual properties. Through emotions, we may gain evaluative knowledge.

But emotions can give us evaluative knowledge only if they are fitting. If an emotion mistakenly presents us with a non-relevant evaluative property, there is no way to say that we gain the relevant knowledge. An example regularly used in the literature is one's fear of a toy bear. In this case, if, unbeknownst to me, the bear is a life-sized toy, then my fear of it is unfitting. We may say that the fear is justified, because I do not know that it is a toy. But since a toy bear is actually not fearsome, my fear is not fitting. Thus, my unfitting fear provides me the wrong information about the fearsomeness of the object. This wrong information is by no means a piece of evaluative knowledge. If emotions can ever deliver evaluative knowledge, they have to be fitting.

This point about the mismatch between a fitting emotion and its object can be seen in other types of cases, such as recalcitrant emotions. In these cases, the subject's evaluative judgment coming from the emotion is inconsistent with her evaluative judgment coming from belief. For example, a person can firmly believe that Skydeck is safe and secure but still feel afraid when standing on it. She can be absolutely convinced that the glass floor does not pose any risk, given that she trusts the designers and the builders. However, when standing on it, looking down at the city through the transparent glass, she cannot help but feel that she is falling. She trembles and sweats. In this case, her fear indicates that the situation is fearsome. But her better judgment tells her that is

not the case; she believes that the situation is actually not fearsome. In fact, it is not fearsome. Accordingly, her fear is not fitting. Furthermore, this fear does not give her relevant evaluative knowledge about the fearsomeness. This example shows that for an emotion to be valuable in terms of providing us with evaluative knowledge, it has to be fitting. Thus, understanding the fittingness of emotions is a necessary step to understanding the value of emotions.

So far, I have shown that both psychologists and philosophers acknowledge the instrumental epistemic value of fitting emotions. Emotions provide useful information and motivate appropriate action. Some philosophers also argue for the epistemic value of emotions in a non-instrumental sense. They contend that emotions are indispensable to our mastering evaluative concepts (Brady, 2013; Deonna & Teroni, 2012; Goldie, 2000; Prinz, 2007; Tappolet, 2016). For example, without ever feeling afraid, we do not understand the concept of fearsomeness. This claim about the value of emotions also presupposes that the emotion at issue is fitting. It focuses on the relationship between the emotion and the relevant evaluative property as its intentional object, which is exactly what the notion of fittingness is proposed to characterize. For this reason, we again need to have a better grasp of the idea of fittingness.

1.2.2.3. Philosophical Research: Moral Value

In addition to these epistemic values, emotions also have moral values, specifically, values related to virtues and characters. Instrumentally, not only can certain moral emotions inform us about moral information and thus convey moral knowledge, but also these emotions can help with building character. For example, self-directed moral

emotions such as guilt and shame are modes of self-assessment, which contribute to building our moral integrity (Antonaccio, 2001). Similarly, pride and admiration in the moral realm can ground moral competence, and thus help build character and nourish moral resilience (Taylor, 1985).

Some emotions also have non-instrumental moral value. They are said to display sensitivities that are dispositions to react in specific ways in response to specific situations. Some of these emotional sensitivities are partially constitutive to virtues. For example, one's feeling ashamed about her treating others unfairly is intrinsically valuable because it manifests the virtue of honesty (Deonna & Teroni, 2015b), anger is intrinsically valuable because it is constitutive of a sense of justice (Srinivasan, 2018), anxiety is intrinsically valuable because it is central to the metacognitive capabilities that constitute virtuous agency (Kurth, 2015, 2018), contempt is a part of integrity in the sense that "virtuous agents will love the good and hate the evil" (Bell, 2013). Because virtues are valuable for their own sake, these constitutive emotions are also valuable for their own sake.

As should be clear, for them to play such valuable roles either instrumentally in character building or non-instrumentally in the possession of virtues, the emotions at issue must be fitting ones. For example, if I feel shameful of finishing a marathon, then it is mistaken to say that this shame manifests anything about honesty or any other virtue. The shame is misplaced. Similarly, if I am angered by someone saying "thank you" to me, this only tells us that I have a problematic sense of justice. If my anxiety is like that of Ivan Chervyakov, the protagonist in Chekhov's novel, "The Death of A Government

Clerk”, who is anxious of offending General Brizzhalov by sneezing right on his head, even if the General has expressed that he does not mind it, it does not seem to display my virtuous agency. Finally, if my contempt is aimed at a colleague who just published a great paper, it does not display integrity; indeed, it appears to undermine my integrity. In summary, all of these alleged values of emotions are based on the condition that they are fitting. The suggestion that unfitting emotions might manifest virtues is absurd.² This also shows the importance of understanding the notion of fittingness.

Other than these moral values, emotions also play important roles in the understanding of our moral life. On the one hand, other-regarding emotions such as respect, love, gratitude, and compassion have been at the center of morality. Many ethical theories take emotions to ground not only special obligations arising from personal relations such as friendship and family ties but also general duties toward humanity (Bagnoli, 2011). On the other hand, negative emotions such as resentment and blame are central in our understanding of moral responsibility (Strawson, 1974). Granted, we talk about things such as “loving unconditionally” sometimes, but in the paradigmatic cases of these reactive attitudes, our discussions are indeed based on the assumption that such attitudes fit their objects (and also of appropriate intensity and

² One may observe that there are cases where even unfitting emotions are arguably manifestations of virtues. For example, Bernard Williams discusses a case where a driver kills a child who runs out in front of his car (Williams, 1981). Although the driver did nothing wrong, many people believe that it is fitting for him to feel a kind of regret and unfitting if he did not feel it. In this case, the regret, even though unfitting, does manifest some virtue of the driver. I suspect that in cases like this, the intuition that the regret is “fitting” goes beyond the technical sense of fittingness, which means to capture the relationship between the particular and the formal object of an emotion. It sounds like that it is *morally inappropriate*, albeit fitting, not to feel regret.

duration). Otherwise, it is less plausible to regard these emotions as expressions of normative expectation, indicators of moral approval, or sanctions against moral transgressions.

To sum up, psychologists and philosophers have proposed many arguments for the value of emotions. These claims are intelligible only if we are talking about fitting emotions. Thus, understanding the concept and the conditions of fittingness is significant for us to understand the values of emotions.

1.2.3. Emotion and Reason

Because of this growing consensus on the significance of emotions, we are offered a valuable opportunity to reexamine some of the central claims from our philosophical heritage as well as to appreciate alternative traditions. One standard understanding of the Western philosophical tradition starting from the Ancient Greece is that emotion is opposed to reason. The division between the rational, the spirited, and the appetitive parts of the soul, originated by Plato and developed by Aristotle, has had long lasting influence. Particularly, in this tradition, emotions are often regarded as destructive, misleading, and irrational, and thus knowledge and rationality should govern and control the emotional parts.

However, the work now being done on emotion can help us see the oversimplification of the opposition of reason and emotion. Amy Schmitter points out that there is a “Mr. Spock” gloss to the contemporary reception of early modern philosophy (Schmitter, 2016). She argues instead that many famous thinkers from the early modern period rejected the idea that passions are erroneous judgments. These

include even Descartes and Spinoza, despite the fact that they were greatly influenced by Stoicism, which advocated the elimination of all emotions (Schmitter, 2005). In contrast, Descartes emphasized the functionality of passions both for practical and theoretical reasoning. His account of wonder is one of the most important early works on “epistemic emotions,” emotions that drive our inquiry. Moreover, Descartes, Spinoza, Malebranche, and Hobbes, as well as sentimentalists such as Hutcheson and Hume all maintain that emotions can be trained and refined to contribute to morality and the good life. In this sense, new research on emotions enables us to “rediscover” these philosophers. As I discussed in the previous section, emotions can contribute to morality and good life only if they are fitting. Thus, understanding the notion of fittingness and the value of fitting emotions allows us to reconsider the works of these important philosophical figures.

This point can be most obviously seen in the case of Kant, who is usually presented as an advocate of reason and an opponent to emotion. However, although Kant’s ethics is centered by reason, some scholars argue that he does not hold that human beings can make appropriate moral judgments in the absence of emotional sensitivity (Wood, 2008). For example, without emotional sensitivity, one may be unable to determine whether a particular act of beneficence is more condescending than kind (*Metaphysics of Morals*, 6: 453). Additionally, Kant emphasizes the importance of respect in his ethics. Our rational grasp of the moral law creates a feeling of reverence, which moves the will to act morally. Accordingly, when one violates the moral law, one feels humiliation; when one acts out of moral law, one feels invigorated and ennobled (Wilson & Denis, 2018). Finally, Kant holds that “it is...an indirect duty to cultivate the

natural feelings in us, and to make use of them as so many means to sympathy based on moral principles and the feeling appropriate to them” (*Metaphysics of Morals*, 6: 457). This emphasis on sympathy is one point among others that makes Kant’s theory closer to Hume’s than is usually recognized (Guyer, 2008).

These examples show that the understanding of fitting emotion partly creates the possibility to “rediscover” these philosophers and make sense of their work. If emotion is purely taken to be contrasted with reason, we might have the tendency to take the “Mr. Spock” view for granted and would not be motivated to investigate whether the view is true or not. Furthermore, it is also hard to make sense of the fact that these philosophers seem to coherently emphasize both emotion and reason. With an understanding of the fittingness of emotions in hand, especially the insight that emotion’s intentionality provides evaluative knowledge, we can see that emotions play a big role in practical reasoning. Thus, we can appreciate the effort of these philosophers to incorporate emotion as well as reason in their philosophical accounts.

Similarly, because we no longer hold that emotion is opposed to reason in a negative way, we can begin to appreciate theories in non-Western traditions and non-traditional thinkers who cherish the role of emotions. For example, Confucian ethics emphasizes cultivating appropriate emotional responses as an important part of being moral, providing an interesting analogy to Aristotelian ethics. Also, understanding the role of emotions helps with apprehending the feminist idea that opposes the false dichotomy of men being rational and women being emotional. Emotions are not simply an absence of reason. Being important in knowledge acquisition and moral action,

emotions are valuable assets that all human beings possess. Again, these discussions are plausible only if we are talking about fitting emotions. Thus, it is a crucial task to understand the notion of fittingness.

1.2.4. Explaining Value

Lastly, fittingness has been a useful construct in theories of the nature of value. Scholars propose to account for value in terms of fitting attitudes (known as the Fitting Attitude analysis of value). For example, someone is admirable just in case one's admiration towards her is fitting. One important reason for this proposal is that it constitutes a middle ground between value subjectivism and robust realism (Deonna & Teroni, 2012; Jacobson, 2011; McHugh & Way, 2016). Value subjectivism is the view that whether something has certain value wholly depends on the subject's actual or dispositional reaction. Robust realism is the view that whether something has certain value depends in crucial ways on some mind-independent properties. Fitting Attitude theories reject both claims. In order to develop this theory of value, however, the key question is how to understand this concept of fittingness. The hope is that once we have a good grasp of this concept, we will solve the hard question of what value is.

1.3. What Is Fittingness?

Now that we have clarified the importance of the notion of fittingness, it is worth mentioning some proposals for how to conceptualize it. There are currently two trends. The reductive accounts include: fitting as having sufficient reasons (Rønnow-Rasmussen, 2011; Rowland, 2013; Scanlon, 1998; Schroeder, 2010; Skorupski, 2007); fitting as good (Hurka, 2001; McHugh, 2012; Moore, 1903; Sylvan, 2012); and fitting as

correct, true, or accurate representation (D'Arms & Jacobson, 2000a; Rosen, 2015; Tappolet, 2016).³ The non-reductive account is that fittingness is just satisfaction of certain standard of correctness; it is normatively fundamental and irreducible to other normative concepts. (Chappell, 2012; Gert, 2016; McHugh & Way, 2016; Schroeder, 2010; Sharadin, 2016). Every account has its pros and cons.

1.3.1. Reasons-based Accounts

The reasons-based account holds that what it is for an attitude to be fitting just is for there to be a sufficient reason to hold the attitude. This approach is structurally similar to the “buck-passing” account of value, which says that what it is for something to be valuable just is for there to be a sufficient reason to value it. The buck-passing account of value analyzes value in terms of reasons. Similarly, the reasons-based version of Fitting Analysis theory analyzes value in terms of fittingness and then analyzes fittingness in terms of reasons. We may say that the latter “passes the buck twice” (Gerken & Kieseewetter, 2017). Thus, although they are different projects, arguments made in the buck-passing account of value are usually used in a parallel fashion for the reasons-based account of fittingness. Both accounts are in line with the “reasons first” approach, which is a well-established approach in the study of normativity (Parfit, 2011; Scanlon, 1998, Skorupski, 2010; Schroeder, 2007). In other words, reason is the most

³ Note that all these reductive accounts try to reduce fittingness to other *normative* concepts. They do not consider the possibility of reducing it to non-normative concepts. I comment on this latter position in the concluding chapter.

fundamental normative concept that can explain other normative concepts such as rationality, ought, value, requirement, permissibility, etc.

An advantage of this reasons-based account is that it provides a simple and elegant explanation of a plausible connection between *fittingness* and *reasons* (Howard, 2018). This can be called the “linking principle.” In the literature, it is widely held that what makes an attitude fitting provides a reason for the attitude. For example, your intelligence makes you admirable and thus fitting to admire; furthermore, your intelligence provides a reason to admire you. This link between fittingness and reasons needs explanation. The reasons-based account has a ready one: if for an attitude to be fitting just is for there to be a sufficient reason to hold it, then it follows directly that what makes an attitude fitting also provides a reason to hold it. As mentioned earlier, in the buck-passing account of value, there is a similar linking principle between *value* and *reason*; it is widely held that what makes something valuable provides a reason to value it. This connection needs an explanation, and the buck-passers have a handy one: if for something to be valuable just is for there to be a sufficient reason to value it, then it follows directly that what makes an object valuable provides a reason to value it (McHugh & Way, 2016; Way, 2013).⁴ As we can see, in providing an account of either

⁴ It may be more straightforward to see this advantage by comparing the reason-first account with the value-first account. According to Thomas Scanlon, the fact that something is good is not an additional reason to value that thing, over and above the non-evaluative facts which constitute reasons to value that thing (Scanlon, 1988, 97). Thus, the proponents of value-first approach cannot explain the connection between reason and value.

fittingness or value, the reasons-based account is using reason as the most fundamental normative concept.

Despite these advantages, there is a serious problem for this account. It is that, although whatever makes an attitude fitting provides a reason for holding it, it is not the case that whatever provides a reason for holding an attitude makes the attitude fitting. In other words, there can be wrong kinds of reasons. Thus, the problem is called the “Wrong Kind of Reasons problem” (D’Arms & Jacobson, 2000a; Rabinowicz & Rønnow-Rasmussen, 2004). We have seen this in the “evil demon” example, where an evil demon asks one to admire him, otherwise he will torture one’s family. One has a reason to admire the demon. But intuitively, the demon is not admirable and admiration towards him is not fitting. Thus, the reason to admire him is the wrong kind of reason with regard to the questions, “Is admiration towards the demon fitting?” and “Is the demon admirable?” A similar message is conveyed by the examples of “immoral funny joke” and “enviable tenured colleague.”

In response, proponents of the reasons-based account have two options. First, they can deny that the wrong kinds of reasons are reasons for feeling the emotions in question at all (Wrong Kind of Reasons skeptics). To illustrate, in the example of the evil demon, it is argued that you only have a reason to *want to* admire the demon, not a reason to admire him (Skorupski, 2010; Way, 2012). However, this response is criticized because, if reason is taken to be the most fundamental normative concept that is irreducible to other concepts, the concept is too thin to make this distinction (McHugh & Way, 2016).

The second response from the reasons-based account is to admit that the wrong kinds of reasons are actually reasons and explain why they are the wrong kinds by specifying conditions for the distinction. Distinctions have been proposed between: object- and state-given reasons; derivative and non-derivative reasons; consequence-dependent and -independent reasons; correctness-independent and correctness-related reasons; and idiosyncratic and shared reasons. None of these proposals has gained wide acceptance. It is argued that they are either circular, or subject to counterexamples, or too complicated to be attractive.

1.3.2. Value-based Accounts

The second account of fittingness depicts it as a kind of goodness. This is usually regarded as in line with the Moorean value-first approach (Dancy, 2000; Hooker & Stratton-Lake, 2005; Scanlon, 1998). There are two main options for this approach, depending on whether goodness is to be understood as predicative (Hurka, 2001; Moore, 1903; Sylvan, 2012) or attributive (McHugh, 2012).

Predicative goodness is also called good simpliciter. Something can be predicatively good in the instrumental or non-instrumental sense. The account is immediately implausible if fittingness means instrumental goodness. For example, in the evil demon example, it is instrumentally good to admire the demon. However, it is not fitting to admire him and he is not admirable. On the other hand, fittingness as non-instrumental goodness is also problematic. It may be true that in examples such as admirableness, someone is admirable just in case it is good for its own sake to admire her. But it is strange to say that someone is enviable just in case it is good for its own

sake to envy her. This problem seems to generalize to many negative emotions and negative evaluative properties.

Something is attributively good when it is a good example of the kind of thing it is. For example, we can make sense of claims such as “It is a good knife” or “It is a good throw (in basketball).” Analogously, it is proposed that we can also make sense of a good attitude in this way. For example, perhaps we can say that the fear one feels towards a bear is a good fear. One obvious problem of this account is that the evaluative property of the object and the attributive value of the attitude can be separated and mismatched. For example, suppose that I admire a person. Suppose further that unbeknownst to me she actually lacks all the features worthy of admiration. In this case, my attitude of admiration can still be a good example of admiration qua admiration. But the object of my admiration lacks admirableness and thus is not fitting to admire (Howard, 2018).

1.3.3. Correct Representation

The third account of fittingness explains it in terms of correct representation. It says that an attitude is fitting just when the attitude correctly represents the object. According to this account, we can say that the joke is funny because it is true that the joke is funny, that your colleague is enviable because it is true that your colleague is enviable, and that the demon is not admirable because it is true that the demon is not admirable.

The most important motivation for this account comes from the analogy between attitudes and beliefs or perceptions. For example, D’Arms and Jacobson regard the relation between an emotion and its object as “a relation analogous to that between a true

belief and the world” (D’Arms & Jacobson, 2000a, 68). Similarly, Tappolet regards emotion as a form of perception (Tappolet, 2016). Just as a perception is fitting when it accurately represents the perceptual property, an emotion is fitting when it actually represents the evaluative property.

However, the biggest problem with this account is that it is circular. Usually, proponents of this account are trying to give a theory of evaluative properties through attitudes, e.g., to explain properties such as funniness in terms of fitting amusement, enviableness in terms of fitting envy, and admirableness in terms of fitting admiration. If fittingness is understood as correct representation, then it leads to the following circular results: something is funny (enviable, admirable) because it is true that it is funny (enviable, admirable).

1.3.4 Fittingness First

Partly motivated by the difficulty of explaining the notion of fittingness in terms of other concepts, some philosophers propose to take the reverse approach to treat fittingness as the most fundamental normative concept and explain others in terms of it (Chappell, 2012; Gert, 2016; Howard, 2018, 2019; McHugh & Way, 2016; Sharadin, 2016). The proponents argue that this is the only way to avoid the many problems raised by the alternative. The debate is still ongoing.

Other than this motivation, recent developments indicate that the discussion about “who’s first” has proceeded on its own (Wodak, 2020). Specifically, this question asks which normative concept (reason, value, or fittingness) is the most fundamental. It is an important question to ask for at least two reasons. First, it helps us adjudicate

between consequentialists and deontologists at the level of normative ethics (Chappell 2012). Second, many philosophers have realized that there are fittingness constraints for beliefs, attitudes (including intentions and emotions), and actions. If there is one most fundamental normative concept, this will give us more expressive power (Gertken & Kiesewetter, 2017). We may say that a drive to unify normativity lies behind this discussion.

However, it is worth noting that although the question of “who’s first” is important, it bears more on the nature of normativity and less on our central task of what fittingness is or how it is distinct from other kinds of normative assessment.⁵

1.4. Summary of the Chapters

My dissertation proposes three constraints that an account of fittingness should meet. Each of the following three chapters is devoted to establishing one of these constraints.

In chapter 2, I uncover a universalist assumption made in the standard account of fittingness. The standard account holds that an emotion is fitting just in case its particular object really instantiates its formal object. As a reminder, the particular object of an emotion is the object whose existence is independent from the subject experiencing the emotion; and the formal object of an emotion is the evaluative property that the subject attributes to the world. I argue that by taking the analogy between emotions and beliefs/perceptions too literally, the standard account assumes that there is always an objectively correct answer to whether the subject has attributed the appropriate

⁵ Also, not everyone accepts that there is an answer to the question, “who’s first?” For example, McDowell (1985) and Wiggins (1987) advocate a “no priority” view.

evaluative property to the world. In other words, the standard account suggests the fittingness condition of emotions is universally applicable to everybody in the same sort of situation. I argue that this assumption is highly controversial by describing several cases where the fittingness criterion of emotions appears to be legitimately disputable across cultures, social classes, and ages. The apparent variation exists not only in cognitively higher and culturally shaped emotions such as admiration, regret, or guilt, but also in evolutionarily basic emotions such as disgust and fear. These cases constitute *prima facie* reasons for questioning the assumption of universal applicability of fittingness standards. More importantly, this shows that any account that assumes such universal applicability of fittingness standards pre-judges controversial metaethical issues about the reality and objectivity of value properties. A better account of fittingness is one that is neutral with respect to such metaethical debates between those who believe that (at least many) values are real and objective and those who believe values are relative either to cultures or to individual valuers.

In chapter 3, I argue that although fittingness assessment is orthogonal to moral justification, as is generally agreed by philosophers of emotion, the judgment of whether an emotion is fitting sometimes is constrained by background moral considerations. I show this point by engaging with the Wrong Kind of Reasons problem in the case of feeling amused by an immoral joke. As has been pointed out earlier, the immorality of a joke is the wrong kind of reason for assessing whether the joke is funny or whether one's amusement towards it is fitting. I distinguish "intramural" from "extramural" moral reasons. Intramural moral reasons are what one has considering the moral feature of the

object. Extramural moral reasons are what one has considering the moral consequence likely to be attached to one's reaction to the object. Based on this distinction, I argue that people judging a joke as not funny for extramural reasons are employing the wrong kind of reasons; but if they do so for intramural reasons, then they are using the right kind of reasons. I also distinguish "not funny" as indicating a negative emotional state with high arousal from "not funny" as indicating a neutral state. I then point out that when people have an intramural reason not to feel amused upon perceiving the moral features of the joke (as opposed to when they consider the consequences of their reaction to the joke), they are in a negative emotional state with high arousal such as moral disgust or anger. Such a negative emotional state is psychologically incompatible with amusement. According to reason internalism and the principle of "ought implies can," if one cannot feel amused, one does not have a reason to feel that way. Thus, one's intramural moral reason not to feel amused implies that one does not have a reason to feel amused. The fact that one does not have a reason to feel amused legitimately allows one to judge that the joke is not funny. By this maneuver, I show that moral reasons can be the right kind of reasons for one's fitting attitude (even if the attitude, i.e., amusement, is not a moral attitude).⁶

In chapter 4, I warn against the tendency to treat fittingness as bearing a value that has been attributed to it—appreciation. I take up this issue by addressing the literature on anger and (affective) appreciation. Amia Srinivasan argues that fitting anger

⁶ The reason to emphasize this point is that for moral attitudes, such as shame, guilt, or moral anger, it is obvious that moral reasons can be the right kind of reasons.

is valuable because through it we affectively appreciate injustice (Srinivasan, 2018). I develop a notion of appreciation. I argue that it is an epistemic emotion that is more than propositional knowledge and is akin to epistemic understanding. In understanding, the subject has more cognitive control over the object and thus forms a deeper relationship with it. Given this account of appreciation, I show that it is an epistemic emotion that is separable from fitting anger. Thus, while it is true that affective appreciation of injustice is valuable, the corresponding fittingness of anger is not valuable in the same way. Fittingness merely involves a standard of correctness, i.e., some correspondence between the attitude and the object. Thus, the value of affective appreciation goes beyond the emotion's being fitting.

1.5. Conclusion

Emotional fittingness is a common evaluation. There are widespread practices of criticizing or confirming them. Emotional fittingness is also an important concept. Vast discussions of the value of emotions are based on it. Understanding this concept can also help us rethink some historical figures' view about the relationship between emotion and reason as well as expose us to alternative philosophical traditions. Finally, philosophers seem to believe that it is a useful notion to solve many difficult questions in metaethics. Given its importance, it is worth thinking carefully how best to characterize it. With my dissertation, I hope to show that there are important constraints in constructing the idea of fittingness.

CHAPTER II

AGAINST THE UNIVERSALIST ASSUMPTION OF EMOTIONAL FITTINGNESS

2.1. Introduction

Emotions are our responses to things or events. I feel afraid of a snake. You feel happy about seeing your best friend. She feels angry about a sexist remark. If an emotion is somehow misdirected towards its object, we may say this emotion is inappropriate. For example, Aiden feels afraid of a bear which he later discovers is a large toy. Intuitively, it seems natural to criticize his fear, given that the toy bear is not worth of fear. In the literature, this kind of normative evaluation is labelled as “fittingness” assessment. To use this terminology, Aiden’s fear towards the toy bear is unfitting.

Fittingness assessment is different from other kinds of normative assessment. An emotion can be *morally* objectionable (for example, feeling amused at a funeral); *prudentially* inappropriate (for example, feeling nervous before giving a public speech); or *aesthetically* unsuitable (for example, feeling proud of poorly made artistic work). An emotion can also be inappropriate in terms of *duration* (for example, holding a grudge for ten years for somebody’s stepping on your foot) or in terms of *intensity* (for example, feeling rage for somebody’s stepping on your foot). These types of assessments are also normative, but they are distinct from the fittingness assessment.⁷ Although in ordinary language, we can call the emotion in each of these cases unfitting, following the

⁷ Though Oded Na’aman argues that fittingness also has to do with the duration. See Na’aman (2020).

literature, I will use “fitting” as a technical term to describe an emotion getting its object right and accordingly use “unfitting” to describe an emotion getting its object wrong.

Scholars have offered various strategies for fleshing out this fittingness criterion. The *standard model* is that an emotion is fitting when its *particular object* really instantiates the *formal object* (D’Arms & Jacobson, 2000b; de Sousa, 1987a; Goldie, 2004; Tappolet, 2016).⁸ In the previous example, the particular object of Aiden’s fear is the toy bear, the formal object is the fearsomeness. According to the standard model, because the toy bear does not really instantiate the property of fearsomeness, Aiden’s fear is not fitting.

In this chapter, I argue that current mainstream characterization of the standard model has made a controversial assumption.⁹ It assumes that the fittingness criterion of an emotion is universally applicable. Call this the universalist assumption. By “universally applicable,” I mean that the criteria of whether a particular emotion (such as admiration) is fitting should not be subject to factors such as who the agent is or what background the agent has. In other words, I argue that the standard model assumes that there is one single absolute answer to whether an emotion is fitting that everybody in the same sort of situation should agree upon. To show that the assumption is highly controversial, I describe in detail several cases where the fittingness criterion of emotions appears to be legitimately disputable. Intuitions about whether certain emotion

⁸ The term “standard model” comes from Scott Howard’s (Howard, 2012).

⁹ In the following, I will use “standard model” as a short term for the “current mainstream characterization of the standard model.” This is because I am not against the standard model in general. As I will comment later, there can be versions of the model where the universal applicability is not assumed.

is fitting vary across cultures, social classes, and ages. The variation exists not only in cognitively higher and culturally shaped emotions such as admiration, regret, or guilt, but also in evolutionarily basic emotions such as disgust or fear. Although there are also cases where people's intuitions easily converge, the cases in which they vary do not appear to be cases where one group of people is mistaken. Instead, the variation seems to show the complexities of human emotional experiences. The breadth, scope, and apparent reasonableness of cases of varied intuitions displays that it is mistaken to simply take for granted that we can work with a universally shared intuition about the fittingness criterion in every scenario. Thus, not only the standard model is problematic, any account that assumes such universal applicability is also problematic.

In section 2, I explain features that motivate the standard model. In section 3, I provide some evidence that the standard model assumes universalism of emotional fittingness. In section 4, I give three examples where the intuitions of emotional fittingness vary. In section 5, I explain how they pose a challenge to universalism. I conclude in section 6.

2.2. The Standard Model

The standard model of emotional fittingness states that an emotion is fitting just in case its particular object really instantiates its formal object (D'Arms & Jacobson, 2000b; de Sousa, 1987a; Goldie, 2004; Tappolet, 2016). To understand this model, we need to understand two well-established claims behind it.

First, it has been widely shared that emotions are intentional, which means that our emotions are always *about* something. For example, Aiden's fear is about the toy

bear. This aboutness means that emotions have intentional objects. The object whose existence in the world is independent from the subject is called the “*particular object*” of the emotion. In our example, the toy bear is the particular object of Aiden’s fear.

Second, it is also widely accepted that emotions are evaluative. When we feel an emotion, we evaluate the object in a certain way. For example, when Aiden feels afraid of the toy bear, he evaluates the object as fearsome, dangerous, or threatening. Such evaluative property is called the “*formal object*” of the emotion. Unlike particular objects, evaluative properties are not independent from the subject but are what the subject attributes to the world. In this sense, they are *response-dependent*. In every case, the emotion as an intentional state has two objects: the particular object and the formal object.

With these two ideas that emotions are intentional and evaluative, we now have a better grasp of the standard model of emotional fittingness. According to this model, whether an emotion is fitting is determined by whether the particular object really instantiates the formal object of the emotion. In our example, the particular object, i.e., the toy bear, does not really instantiate the formal object, i.e., the evaluative property of fearsomeness. Thus, Aiden’s fear is not fitting. In general, the standard model answers the fittingness question by urging us to look at whether an emotion’s formal object is correctly instantiated by the particular object. To use D’Arms and Jacobson’s words: “considerations of fittingness...bear on whether the emotion’s evaluation of the circumstances gets it right: whether the situation really is shameful, funny, fearsome, and so forth” (D’Arms & Jacobson, 2003, 132).

The standard model is attractive for three reasons. First, it retains the intuitive idea that emotions can be normatively assessed with respect to their objects. Second, it has resources to answer to the Wrong Kind of Reasons problem. Third, examples such as Aiden's fear are highly uncontroversial. I will explain each of these points in turn.

The idea that emotions are criticizable in terms of fittingness is intuitive. The proper account of the nature of emotions is still debated among scholars. However, no matter whether one holds that emotions are more like propositional attitudes such as beliefs or beliefs plus desires, more like perceptions, or *sui generis*, scholars take it is a desideratum for their theories to leave room for fittingness assessment. As long as it is possible for one to mistake the object of an emotion, the fittingness question will need to be addressed.

The Wrong Kind of Reasons problem has generated quite a body of literature.¹⁰ The idea is that a type of examples shows that the reason for holding an emotion can be the wrong kind. One such example that has been extensively discussed is the evil demon example. Imagine that an evil demon asks you to admire him, otherwise he will torture your family severely. In this case, you have a reason to admire him. But intuitively, this is the wrong kind of reason for admiration. On the contrary, we would want to say that the evil demon is not admirable. Accordingly, if you admire him, your admiration is not fitting. Using examples such as this, scholars tend to argue that the right kind of reasons for attitudes (or even actions) have to do with the correctness of the attitude. Taking this

¹⁰ For a useful overview, see Gertken & Kiesewetter (2017).

line of thought to the specific context of emotional fittingness, we can see that the right kind of reasons for an emotion has to do with the correctness of the emotion, i.e., whether the evaluative property as the formal object is correctly instantiated by the particular object. Thus, the standard model has resources to reply to the Wrong Kind of Reasons problem.

The third point that makes the standard model attractive is that the intuitions pumped by examples such as Aiden's fear, which is usually used as the introductory example, are highly agreed upon. In judging that Aiden's fear is not fitting, we share the intuition that that we can claim things such as "the toy bear is not *really* or *actually* fearsome." In other words, in examples like this, we are confident that we can get to the *fact* about which object has the corresponding evaluative property; we are also confident that we *correctly* get to the evaluative property.

This third point is crucial. General claims such as "X is *really* or *actually* Y", where X is a particular object and Y an evaluative object, are not as obvious as the examples such as Aiden's fear suggest. I contend that our intuitions vary greatly across different examples. Maybe we find the example of fear of a toy bear compelling, since we share the understanding of fearsomeness in this case. It is hard to imagine anyone would judge a toy bear, knowing that it is merely a toy, to be fearsome. However, it is a mistake to come to a general conclusion that fittingness conditions apply universally based on intuitions about such examples. For one thing, there are more contentious emotional types where people do not agree on when the formal object is instantiated correctly, such as pride or shame. For another thing, our intuitions for even one emotion

type can vary from concrete case to case. Before describing such examples, let me show that the standard model does indeed make the universalist assumption.

2.3. The Universalist Assumption

In this section, I give some evidence for thinking that the standard model assumes that the fittingness condition of emotions is universally applicable to everybody in the same sort of situation. I take the work of Justin D'Arms and Daniel Jacobson and of Christine Tappolet as two examples. D'Arms and Jacobson's work is one of the first that proposes emotional fittingness as a distinct notion of normative assessment. Tappolet's work is representative of the most systemic account of emotional fittingness in the contemporary literature.

2.3.1. D'Arms and Jacobson's Rational Neo-Sentimentalism

In "Sentiment and Value" (2000b), D'Arms and Jacobson argue that the work by Simon Blackburn, Allan Gibbard, John McDowell, and David Wiggins share the central tenet of Neo-Sentimentalism: to think that something has certain evaluative property is to think it appropriate to feel certain emotion in response to it. They further argue that the notion of appropriateness offered by these philosophers fails because they do not distinguish the right kind of reasons from the wrong kind of reasons to hold the attitude. However, D'Arms and Jacobson do not thus reject Neo-Sentimentalism; instead, they sketch their own version while trying to avoid the problem. Their proposal is to "examine our actual emotions piecemeal, in order to articulate differences in how each emotion presents some feature of the world to us when we are in its grip" (746). In specifying this

piecemeal strategy, I argue that D'Arms and Jacobson assume the universal applicability of fittingness criterion.

My first piece of evidence concerns their terminology. In their critique, they use phrases such as the following (the emphasis is mine):

the response is appropriate in the relevant sense: that the object of one's outrage *really* is outrageous. (735)

We can examine shame, for instance, in order to decide which things are *truly* shameful. (736)

In other words, we are stipulating that to judge F fitting is to endorse the response in the relevant way, which constitutes taking the circumstances to be *genuinely* F. (746)

In the commonsensical use of the language, the terms such as “really”, “truly”, and “genuinely” imply that the correct answer for the question of whether a particular object instantiates certain evaluative property should be shared by everybody, no matter who the subject of the emotion is and what situation she is in. In other words, if two people disagree about whether a subject's emotion is fitting, there is a normative reason to believe that at least one of them is making a mistake.¹¹

Second, according to their explanation, the Neo-Sentimentalist accounts offered by Blackburn, Gibbard, McDowell, and Wiggins all aim to achieve some kind of truth.

¹¹ A non-universalist can also use terms such as “really”, “truly”, or “genuinely”. In this sense, this piece of evidence is not a knock-down argument that the scholars assume universalism. However, as will be clear in the following, this is the best interpretation of their work.

Although criticizing the specific versions offered by those authors, D'Arms and Jacobson follow the suit in assuming that this kind of truth is required in a Neo-Sentimentalist account.

McDowell and Blackburn agree that, given the subjective basis of values in human responses, talk of the *truth* of evaluative judgment must be earned, rather than simply assumed. Some philosophers find the idea of earning talk of truth in these contexts to be obscure or even disreputable and would prefer to speak of earning the right to claim evaluative *knowledge*. Others frame the debate as being over the *objectivity* of judgments that are granted to be founded in subjective human responses. (733, my emphasis)

The criteria of “truth”, “objectivity of evaluative judgments”, or “such judgments manifesting knowledge” have also been explicitly mentioned elsewhere (736, 738). As I understand it, these criteria are essential because they make fittingness a normative concept. They also imply the following: when one claims something is true, one implies that people who do not share that opinion are mistaken, or, when a person claims something is true, she does not mean that the thing is just true *for her*; otherwise, the usage of the term such as “truth” loses its normative force. As far as I see it, this is just another way to say the fittingness criterion is universally applicable.

Similar textual evidence can be found in their other important works (D'Arms & Jacobson, 2000a; D'Arms, 2013). To clarify, I am not saying that they take evaluative properties to be response-independent. As proponents of Neo-Sentimentalism, they clearly deny that. What I contend is that they assume the *universal applicability* of the

fittingness criterion. One may object that the phrases of “really”, “truly”, or “genuinely” and the talk of “truth”, “objectivity of evaluative judgments,” or “such judgments manifesting knowledge” do not necessarily imply that whenever an emotion is fitting along these lines it must be fitting universally. The objection goes, if we work with a *plural* or *relative* notion of truth, it is possible that we can talk about something being true without committing to the claim that it is universally true. The point is that the above-mentioned expressions are compatible with *truth pluralism* or *truth relativism*.

While this objection is legitimate, I do not think truth pluralism or relativism is what D’Arms and Jacobson have in mind. On the contrary, I think they use these criteria as implying the universal applicability of the fittingness criterion. While they do not explicitly make this claim, it is reasonable to interpret that they do. For example, at one point they discuss community standards:

If one’s community deems some trait shameful (one’s sexual preference, say), it will be difficult not to be ashamed of it, but surely one need not take such community standards to establish that it *really* is shameful. Perhaps this is a prejudice of the culture which is not to be acceded to but resisted. (737, my emphasis)

Later, they write, “community standards do not settle whether X is F’ (738, my emphasis).

I agree with these remarks. But I think that D’Arms and Jacobson discuss community standard in order to show an example that fails to establish the truth, objectivity, or

knowledge of evaluative claims rather than to suggest that a normative standard can be plural or relative.

2.3.2. *Tappolet's Representational Neo-Sentimentalism*

Tappolet's Representational Neo-Sentimentalism is a prominent effort to explain the notion of fittingness. According to this theory, something has certain evaluative property if and only if it is such that feeling certain emotion is a *correct* response to it. Here, the notion of correctness implies an epistemic point and fits well with her particular theory of emotion. According to her perceptualist theory, emotions are intentional states that represent evaluative properties in the similar way as perceptions represent perceptual properties. For example, when I see a ball, my eye represents the shape property "sphericity." Similarly, for Tappolet, when I admire Mother Teresa, my admiration represents the evaluative property "admirableness." Like perceptions, emotions also have correctness conditions. My representation of the property is correct if and only if the object *indeed* possesses the property. The notion of "correctness" is what "fittingness" means.

The perceptualist theory of emotion is the most popular one in the current literature. I want to show that this account also assumes the universal applicability of fittingness criterion. First, it is quite natural to see this assumption by thinking about the perception analogy that Tappolet gives. It is plausible to think that the criterion of whether one fittingly or correctly takes something to have certain perceptual property (for example, color or shape property) is universally applicable. For perceptual properties, we usually think that the correctness conditions of them are not and should

not be disputable among subjects. We assume a cross-human criterion of grasping those perceptual properties. Facing the same object, the claim, “it is a ball,” and the claim, “it is not a ball,” cannot be both correct. Because Tappolet takes emotions to be literally a kind of perception, this assumption about perceptual properties must also be shared by her account of emotions. According to her theory, facing the same particular object, the claim, “it is fearsome,” and the claim, “it is not fearsome,” cannot be both correct. Accordingly, the claim, “the fear toward it is fitting,” and the claim, “the fear toward it is not fitting,” cannot be both correct. Thus, the perceptualist theory assumes that the fittingness criterion for an emotion is universally applicable to everyone in the same sort of situation.

Second, we can also see direct textual evidence for this assumption. Tappolet emphasizes that in her account the concept of fittingness is like the following (my emphasis):

[Fittingness] is a matter of representing things *as they are* (Tappolet 2016, 87).

[Emotions] represent their objects as having certain evaluative properties and are correct depending on *whether things are* as they are represented. For instance, an emotion of admiration with respect to a friend will be correct just in case the friend is *really* admirable (87-88).

...to judge that an emotion is correct is to be committed to the claim that its object *really* has the corresponding evaluative property (91).

Like the discussion of D’Arms and Jacobson’s work, in these quotes, phrases such as “things as they are” and “really” imply that the correct answer for the question whether a

particular object has certain evaluative property should be shared by everybody, no matter who the subject who experiences the emotion is and what situation she is in. In other words, if two people disagree about whether a particular experience of an emotion is fitting, there is a normative reason to believe that at least one is making a mistake.¹²

Similar expressions showing the universalist assumption can be found in other perceptualist accounts of emotions (de Sousa, 1987a; Goldie, 2004). The issue of what an emotion is and the issue of when an emotion is fitting are distinct. But one's theory of emotion can easily make one have certain commitments about its fittingness. The perceptualist theory is one of the most popular contemporary theories of emotions. No matter whether a perceptualist takes emotions to be literally a kind of perception, like Tappolet, or only takes perception to be a useful analogy, it seems natural that they assume a universally applicable criterion of fittingness. After all, that is the case for perceptions.¹³

In sum, the two influential and representative accounts of emotional fittingness offered by D'Arms and Jacobson and by Tappolet both assume that the fittingness criterion of an emotion is universally applicable, as in the case of the correctness criteria

¹² Again, these terms themselves do not constitute a knock-down argument that she assumes universalism. But this is the best interpretation of her work.

¹³ Arguably, Jesse Prinz's perceptualist theory is an exception (Prinz, 2004, 2007). Prinz has a complicated perceptualist-somatic theory. In his theory, worldly objects cause bodily changes. Emotions are just perceptions of such bodily changes. Meanwhile, emotions represent not bodily changes but evaluative properties. Whether a particular worldly object instantiates a certain evaluative property is determined by one's calibration file. Because technically speaking, it is possible for people to develop different calibration files, the question of correct evaluative property instantiation can be subject-relative. Thus, although Prinz's theory is perceptualist, he does not have to commit to the assumption of universal applicability of emotional fittingness.

of beliefs and perceptions. These accounts presuppose that the normative sense of fittingness means that there is no room for disagreements about when a particular object instantiates an evaluative property and thus when an emotion is fitting. Everyone in similar circumstances would fittingly have the same emotion.

This assumption is implicitly taken for granted in the literature for two reasons: problematic generalization or problematic analogy. First, scholars usually introduce the question using “easy” examples like Aiden’s fear of a toy bear, where it seems uncontroversial that the toy bear does not really instantiate the property of fearsomeness. Then they implicitly generalize this shared intuition about correct property instantiation to all cases, i.e., in all cases, we have the same agreed intuition about whether a particular object really instantiates a formal object. This implicit generalization removes vast number of “difficult” cases where our intuitions about whether a particular object really instantiates a formal object legitimately vary.

Second, as shown above, scholars usually use analogies from belief or perception to discuss emotional fittingness. It is fairly uncontroversial that the fittingness criterion of belief, i.e., truth, and the fittingness criterion of shape perception, i.e., correctness, are universally applicable. Using these analogies, scholars seem to take it as a working hypothesis without specific argument that the fittingness criterion of emotion, whatever it is, is also universally applicable. In the next section, I give some examples that they pose challenges to the universalist assumption.

2.4. Cases of Varied Intuitions

In this section, I offer some examples to show that culture, social status, and age can influence our intuitions about fittingness and evaluative property instantiation. These examples all share the feature that there can be contrasting reactions about emotional fittingness that each of which seem reasonable.

Culture

I believe that everyone can make sense of the phenomenon of admiring somebody. I invite the reader to use their intuitions to assess whether the admiration discussed in the following scenario is fitting or not. A person was seriously ill and would die soon. At his deathbed, he says to his friends, “Look at me. My whole body is intact. My hands and my feet are all here. I did not lose any body part. I have spent my entire life to protect it, the body that my parents gave to me. Now I am dying. But I have fulfilled the duty of filial piety to my parents.” His friends think that by maintaining his body intact he truly fulfilled the duty of filial piety. Thus, they truly admire his effort. Is this admiration fitting?

At least a considerable number of people, for example, most contemporary Chinese, will judge that the admiration is not fitting. They probably would think maintaining one’s body intact in order to fulfill one’s duty of filial piety does not instantiate the property of admirableness. According to the standard model, this implies that the admiration is not fitting.

I personally share this intuition that the admiration is unfitting. However, I am also aware of people who might not share it. Actually, if I were to live 2000 years ago in

China, I probably would not share it myself. In fact, the scenario I described was recorded in *The Analects* (8.3). The dying person is a student of Confucius. His name is Zengzi, who was famous for being a model of filial piety and thus one of the moral saints. Of course, *The Analects* does not say that his disciples (who are his friends in my modified example) admired him for what he said. But it is not a stretch to say so. *The Analects* was regarded as the most important material for moral education. In that period of time, to cultivate a moral character, admiring a role model was considered necessary. In other words, feeling admiration towards people who maintain their body intact in order to fulfill the duty of filial piety is exactly the correct emotion to cultivate. If a person does not feel admiration, it is that person who has the wrong attitude. Thus, if I were to live 2000 years ago in China, I not only *would not* share the intuition that the admiration is unfitting but also *should not* do so.

I have given the example of admiration. It may be objected that admiration is an emotion that requires higher cognitive capacity but more basic emotions do not have this requirement. Basic emotions include fear, joy, sadness, anger, disgust, and contempt. There are some disagreements on what should be included in this list. But the idea is that these emotions are universal and evolutionarily early. Thus, the objection continues, there are a vast number of cases where our intuitions of fittingness are shared. In the next two examples, I reply to this objection. First, I consider disgust, an emotion that is uncontroversially basic. Using this example, I show that variation in intuitions about the fittingness of basic emotions is also common. Second, I discuss a set of higher emotions,

nostalgia, regret, claustrophobia, emptiness, and fear.¹⁴ Using this set of examples, I show that admiration is by no means one of the few higher emotions that we experience differently. Many significant emotions are complex and cognitively demanding. In these cases, disagreement about fittingness is also common. Thus, it would be a serious flaw if the standard model leaves such emotions out of its explanation framework.

Social Status

The 2019 Oscar winner is the South Korean movie *Parasite*. It tells the story of a poor family (a father, a mother, a son, and a daughter), living in a semi-basement apartment and struggling financially, who manage to get hired by a wealthy family, the Park family. First, the son is successfully recommended by his friend to be an English tutor for the family. Then, the poor family uses fraud to get the daughter hired as an artistic counselor, the father as a driver, and finally the mother as a housekeeper. Although both families seem happy with the changes in their lives, Mr. Park has long noticed a bad smell strangely surrounding him ever since the new hiring, without understanding where it comes from. As the audience, we know that it is the smell of the semi-basement coming from the driver; the movie has one scene where the poor family discusses the exact smell. Towards the end of the movie, Mr. Park cannot help reacting to the smell by covering his nose. This act is regarded as extremely insulting by the driver. Out of rage, the driver kills Mr. Park. Now, is Mr. Park's emotion of disgust fitting?

¹⁴ As it will become clear, the “fear” here refers to the emotion when one thinks back of her past and look forward to her future, not simply fear of a physically threatening object. It is in this sense that fear in this case is cognitively more demanding.

If you have seen the movie, you will definitely believe that Mr. Park sincerely feels disgusted by the smell. He notices it at the moment when he has to take every minute possible to save his son. Yet he still reacts to the smell by covering his nose. The muscle on his face is contracting due to the disgust he feels. I think we can only understand this reaction as a strong automatic one; otherwise, he would have taken the time, however short, to save his son. This strong automatic reaction nicely captures the relevant feature of disgust: it is a basic emotion.

But is this sincerely felt emotion fitting? Mr. Park probably would say yes. Presumably, it seems the musty smell composed of mold, urine, feces, dead mice, or sewer gas indeed really instantiates the property of disgustingness. This would also be the judgment from the point of view of most observers: Mr. Park's disgust is fitting.

Would the driver agree? Granted, he is not irritated because he disagrees with Mr. Park about whether the smell is disgusting. Rather, he is motivated to kill because he regards Mr. Park's disgusted reaction as seriously insulting. Still, if we ask him, "Do you feel disgusted by the musty smell of your apartment that is composed of mold, urine, feces, dead mice, and sewer gas?", it would not be surprising if he sincerely says no. It is the place he has to stay every night. He cannot afford to feel disgusted by it. Actually, in one scene of the movie, he smells his clothes very hard, but seems not to be bothered by it at all.¹⁵ Thus, even after being told by his daughter that it is the smell of the semi-

¹⁵ It is possible that he loses the ability to register the smell, due to his long time living in the basement. If this is the case, then I cannot say that he is not disgusted by the smell, after all, he does not experience the smell in the first place. But I think it is also possible that even though he has been living in the basement for a long time, his olfactory capability is still not impaired. For example, it is

basement, and knowing all very well where it comes from, he still does not feel disgusted. He may just hold that the musty smell of his apartment that is composed of mold, urine, feces, dead mice, and sewer gas does not instantiate the property of disgustingness.

Age

The last factor that influences our intuitions about emotional fittingness that I would like to discuss is age. Age brings about different personal experience. Age also brings about corresponding bodily changes. These factors working together can make one's evaluation and emotional responses radically different, not only interpersonally but also intrapersonally.¹⁶ The concrete example I am going to discuss on this point is midlife crisis.

In his nice short book *Midlife: A Philosophical Guide*, Kieran Setiya describes what his feeling of midlife crisis is like.

I started thinking about midlife about six years ago, at the tender age of thirty-five. On the surface, life was going well. I had a stable family and career. I was a tenured professor in a good department housed in a congenial Midwestern city. I knew I was lucky to be doing what I loved. And yet there was something hollow in the prospect of doing more of it, in the projected sequence of accomplishments stretching through the future to retirement, decline, and death. When I paused to contemplate the life I had worked so hard to build, I felt a disconcerting mixture of *nostalgia, regret, claustrophobia, emptiness, and fear*. Was I having a midlife crisis? (Setiya, 2017, 2, my emphasis)

possible that although he cannot recognize the smell coming from his clothes, he can recognize the same smell coming from other sources. If so, then he perceives the smell and is not bothered by it. I speak to this possibility.

¹⁶ Here it is good to mention that scholars widely agree that emotions have an important physiological dimension. Thus, it seems even more adequate to bear this in mind in the discussion of fittingness of emotions.

For our purposes, this is a perfect passage to reflect on. Towards the end, we can see clear emotion words such as “nostalgia, regret, claustrophobia, emptiness, and fear.” Of course, such feelings are not directed towards the fact that Setiya was living a decent life as he described in the passage. Rather, throughout the book, we can find that these feelings are about things such as awareness of mortality, the finitude of life, lost opportunities, failed ambitions, etc.

Are these things peculiar to midlife? I don't think so. I think it is highly possible that an 18-year-old Ivy League college student can be aware of mortality and the finitude of life (perhaps through an introduction to philosophy class), can be reminded of the experience of giving up the cherished violin to major in pre-law, can encounter failures such as being rejected by a certain school fellowship program. Meanwhile, that it is called *midlife* crisis indicates that such a crisis is not as likely to occur at the age of 18. I think it should not be astonishing if this student in our scenario says that she does not feel “nostalgia, regret, claustrophobia, emptiness, and fear” at all. Her reaction is perhaps even the normal reaction. By that I mean, we usually think the young should not have these feelings towards what they experienced. After all, they are young and hopeful. Thus, for the young, the lack of such feelings seem fitting.

Certainly, we may think that such awareness, lost opportunities, and failed ambitions experienced in the young college student are nothing in comparison with those in midlife, when reactions to those things are both deeper and wider in scope, and more serious in effect. After all, in comparison, a person in her midlife is indeed nearer to death, indeed has fewer opportunities, and indeed finds it more difficult to change

career. However, if you find the comparison with young people unconvincing, now think about an elderly person. Surely the elderly can experience awareness of mortality, the finitude of life, lost opportunities, and failed ambitions. In this respect, whatever fact is true for a person at mid-age perhaps is also true for an elderly person. After all, in comparison with the mid-aged, the elderly person is indeed nearer to death, indeed has fewer opportunities, and indeed finds it more difficult to change career, if there is any chance of a career at all. But do the elderly feel “nostalgia, regret, claustrophobia, emptiness, and fear”? According to Setiya, it seems not as frequent. Life satisfaction seems to rise at old age. As I understand it, even though the fact, for example, that a philosopher can no longer change her career to be a musician, she evaluates and feels towards this fact differently at different age. At middle age, she may feel nostalgia or regret. At an older age, she may not feel that way. To use our terminology, the particular object of the emotion is the same, but this particular object instantiates the relevant evaluative property and thus becomes the formal object of an emotion at one moment while not instantiating it at another moment. Now, the question for the standard model is: Which is the correct instantiation? The one at middle age, or at older age?

I have discussed several examples. Let me summarize the main points. Using the example of admiration, a non-basic emotion, I showed that the intuitions regarding the fittingness of this emotion vary in different cultures. Using the example of disgust, a basic emotion, I showed that intuitions of fittingness vary in people with different social status. Using the example of the bundle of “nostalgia, regret, claustrophobia, emptiness, and fear”, I showed that intuitions of fittingness vary in people of different ages. I have

included basic and non-basic emotions. I have included the interpersonal and the intrapersonal dimension. I have discussed factors that may affect one's intuition of fittingness such as culture, social status, and age. I am sure there are other factors. Yet this is already a complex picture. In this picture, we have dozens of emotions. We have two dimensions. We have three varying factors. The fittingness of any one of the dozens of emotions can be discussed along one of the two dimensions taking any of the three varying factors into consideration. The point is, all the examples I discussed show that there can be intuitively fitting emotional reactions that are not universally applicable.

2.5. What to Make of the Disagreement?

Disagreements themselves do not show the falsity of universalism. This can be seen most clearly in the metaethical discussion on moral relativism and moral realism. One central claim of moral realism is that the truth or falsity of moral judgments is absolute or universal.¹⁷ Moral relativists object to this claim. One strategy that a moral relativist commonly uses is the argument from moral disagreement (Harman, 1996; Prinz, 2007; Wong, 2006). According to this argument, the fact that there are fundamental moral disagreements poses a serious challenge to moral realism. However, to use the disagreements to further undermine realism and to support relativism, more premises are needed. For example, the fact that two scientists can disagree on whether string theory is

¹⁷ People use the term “moral realism” in different ways. In one use, moral realism is the view that there are objective moral facts. In this sense, moral relativists are not necessarily anti-realists because relativists can also hold that there are objective moral facts. They only disagree with realists on whether the facts exist universally or not. To put it another way, moral realism in this sense and moral relativism both belong to moral objectivism. In another use, moral realism is the view that moral facts exist absolutely or universally. In this sense, moral relativists are anti-realists. See Joyce (2016). It is the second sense of moral realism that is in use here.

true does not imply that there is not a universal, objective fact of the matter to the question. Both relativists and realists must say more about how to explain these disagreements.

There are a couple of arguments each side can make. First, a realist can say that the disagreements are only about non-moral facts and “careful philosophical examination will reveal...that agreement on nonmoral issues would eliminate almost all disagreement about the sorts of moral issues which arise in ordinary moral practice” (Boyd, 1988, 213; also Brink, 1984). The idea is that disagreements about *non*-moral facts do not undermine *moral* realism. In response, a relativist must show that even though all the non-moral facts are agreed upon, there still will exist disagreements on moral matters.

Second, a realist can say that “disagreement suggests a fault of at least one of the interlocutors...some irrational emotional response that stands as a barrier to moral convergence” (Shafer-Landau, 1994; 331). Thus, in ideal scenario where people are fully rational and well informed, there will be no moral disagreements. In response, a relativist must show that even in the ideal scenario, the moral disagreements will still not be resolved.

Third, a realist can say that even if no moral convergence is achieved, moral realism is still not threatened. One line of this argument is to say that moral reasoning involves self-evident moral beliefs, i.e., beliefs about things that “adequately understanding and attentively considering them is sufficient to justify believing them” (Shafer-Landau, 2003, 248). In response, a relativist must show how moral disagreement

can indeed undermine self-evident moral beliefs.¹⁸ Another line of this argument is to say that a moral realist only needs to maintain that *most* genuine moral disputes are resolvable because certain types of irresolvable moral disagreements do not threaten moral realism (Brink, 1984, 424). In response, a relativist must show how the existence of some moral disagreements show the incommensurability of values and thus threaten moral realism.¹⁹

Finally, a realist can say that the plausibility of moral realism does not depend on the assumption that we have direct, non-inferential access to morality.²⁰ The idea is, what is true may not be *knowable* to reasonable people. In response, a relativist must show how the idea of unknowable moral truth is implausible.

Each side of the debate has to explain how their treatment of the phenomenon of moral disagreement is more plausible than the other side. But no matter what, it seems to be the consensus that some explanations are needed in response to the phenomenon. The existence of moral disagreement constitutes at least an apparent challenge to moral realism.

Related to emotional fittingness, all I want to show is that the examples I gave pose a *prima facie* challenge to universalism, just like the phenomenon of moral disagreement poses a *prima facie* challenge to moral realism. Although the fact that it is a challenge does not immediately imply that universalism is wrong, it suggests that

¹⁸ For example, Sinnott-Armstrong argues that even though some beliefs are self-evident, *maintaining* those beliefs when being challenged may not be self-evident. As he argues, one needs to do more than just repeat oneself in justifying her belief in the face of disagreement (2002).

¹⁹ Note that realism here implies universalism.

²⁰ Loeb (1998) mentions this line of argument.

universalists face the task of giving some explanations about these cases, just like moral realists face the task of giving some explanations about the phenomenon of moral disagreement. Universalists may have a good explanation, perhaps along the lines invoked by moral realists sketched above, but they should not ignore such cases and assume universalism. It is one question whether these examples make reasonable cases that challenge universalism, it is a further question whether universalists can successfully meet this challenge. It is the first question that I am answering in this chapter.

There are two clarifications needed to be made. First, although the examples I gave show apparent disagreements on fittingness judgments, there are obviously cases where there are universally applicable fittingness judgments. For example, perhaps in all cultures, one's reaction towards somebody's torturing one's children is anger and this anger is fitting. This poses a different constraint to an account of fittingness because any such account should be able to explain why in some cases there are indeed universal judgments about fittingness. While universalists have a ready answer, it is now the non-universalists' turn to face the challenge. They will need to explain why looking at the specific context of the emotion would sometimes yield the convergent judgments about fittingness.

The second clarification is that even though I criticize the universalist assumption, I do not object that there is a fact of the matter about whether a particular emotion is fitting. On this point, I agree with the standard model advocates. We can see this point by looking at the relevant metaethical discussion again. Some people believe

that moral judgments have truth value. That is, we can say something like “It is true that murder is morally wrong.” On the contrary, some people believe that moral judgments do not have truth value. That is, statements such as “It is true that murder is morally wrong” is mistaken, not because, according to this view, it is false that murder is morally wrong, but because the judgment that murder is morally wrong is not apt for truth evaluation. Among people who believe that moral judgments are subject to truth evaluation, opinions diverge on how we should evaluate their truth value. One view is that moral judgments are true or false in a universal or absolute sense. The other view is that the truth value of a moral judgment is relative to culture or to the individual doing the judging.

Similarly, we can delineate fittingness judgments in this way. There is the view that fittingness judgments have truth value. That is, we can say something like “It is true that admiration is fitting in this case.” On the contrary, there can be the view that fittingness judgments do not have truth values. That is, statements such as, “It is true that admiration is fitting,” is mistaken, not because it is false that admiration is fitting, but because the judgment that admiration is fitting is not apt for truth evaluation. Although I criticize the universalist assumption, we both belong to the first camp. That is, we both believe that fittingness judgments are subject to truth evaluation. We both regard the question “Is this emotion fitting?” as answerable rather than misguided. While universalists assume that fittingness judgments are true or false in a universal or absolute sense, I challenge this assumption.

These examples show that the more we discuss a concrete example, the more complicated the question of fittingness is. This is perhaps a reflection of the complexity of human emotional experience. It also means that maybe all the concrete cases deserve some detailed discussion. At least this should be done before a conclusion about an absolute and context-insensitive notion of fittingness is drawn. Scholars usually start from easy examples such as Aiden's fear of a toy bear and then try to get something in general about fittingness, as if trusting that in every case the intuition will be widely shared. It is true that we share the intuition that, for example, Aiden's fear is unfitting. But it is problematic to assume that well-informed, reasonable people will always evaluate fittingness in the same way. When such disagreements do arise, as the examples of admiration, disgust, and the set of nostalgia, regret, claustrophobia, emptiness, and fear show, the shakiness of the universalist assumption becomes apparent. Thus, it is problematic to assume a universally applied normative criterion of fittingness without argument. An account of fittingness that is neutral to these meta-ethical issues will be more plausible.

2.6. Conclusion

In this paper, I criticized the standard model of emotional fittingness. The model says that an emotion is fitting if and only if the particular object really instantiates the formal object. The model is attractive because it captures the desideratum that emotions are subject to fittingness assessment and it has resources to reply to the Wrong Kind of Reasons problem. I argued that despite these good features, the model is problematic because it assumes a controversial meta-ethical thesis that the fittingness criterion is

applicable universally to everybody in the same sort of situation. I suggested that this assumption has crept into the literature because scholars rely too heavily on easy examples where our intuitions about fittingness converge to a large extent and because they take analogies with perception too literally. I argued that it is problematic to generalize shared intuitions in easy examples to intuitions in all the emotion cases. I have shown that there are numerous examples in which our intuitions diverge, which calls into question the use of the belief or perception model. Thus, these examples pose a serious challenge to the standard model.

My recommendation is that the standard model must give up its assumption of universal applicability. But even if fittingness is relative to culture, social status, or age, it can still be normative because it is still possible for an emotion to be unfitting. Besides, even if fittingness is relative in some way, different emotions will still be tied to some kinds of reasons rather than others. Thus, giving up the universalist assumption does not prevent a notion of fittingness from being normative and having the resources to solve the Wrong Kind of Reasons problem.

CHAPTER III

JOKES CAN FAIL TO BE FUNNY BECAUSE THEY ARE IMMORAL: THE INCOMPATIBILITY OF EMOTIONS*

3.1. Introduction

In reaction to a morally offensive joke, some people may feel so offended that they do not feel amused (Yam, Barnes, Leavitt, Wei, Lau, & Uhlmann, 2019; Yam, Christian, Wei, Liao, & Nai, 2018). Philosophers of emotion frequently charge these people with committing a fallacy, namely, the mistake of confusing the reason for amusement and the reason for moral disapproval. In this paper, I argue that, at least in some possible cases, people do not commit this fallacy. In other words, immoral jokes can fail to be funny *because* they are immoral.

Before proceeding, we need to distinguish my normative claim from a descriptive claim. The *descriptive* claim is that some people actually don't find a joke funny because they think the joke is immoral. This empirical question is *not* the contention of this paper. Instead, I discuss the *normative* claim, i.e., people can have *legitimate reasons* to consider morality in judging the funniness of a joke, or can be *justified* in making that judgment.²¹ Thus, the normative claim does not depend on the

* Reprinted with permission from “Jokes can fail to be funny because they are immoral: The incompatibility of emotions” by Dong An & Kaiyuan Chen, 2021, *Philosophical Psychology*, 34:3, 374-396, <https://doi.org/10.1080/09515089.2021.1874331>. Copyright 2021 by Taylor & Francis Group.

²¹ There is a distinction between “a legitimate reason from the perspective of morality” and “a legitimate reason from the perspective of humor.” It is this latter sense of legitimacy that is in use in my arguments. In this sense, what is concerned is the funniness of a joke rather than the morality of

fact that there is such a group of people. Even if they do not exist, the normative claim (and thereby the argument of this paper) is still sensible.²²

My argument turns on distinguishing two kinds of moral reasons, which I call “intramural” and “extramural.” Intramural moral reasons are what one has considering the moral feature of the object. Extramural moral reasons are what one has considering the moral consequence implied in one’s reaction to the object. Based on this distinction, I argue that, if people judge a joke not to be funny for extramural reasons, they are making the “moralistic fallacy”; but if they do so for intramural reasons then they are not. This is the central claim of this chapter. The moralistic fallacy is the mistake of evaluating the funniness of a joke based on the consideration of whether it is morally appropriate to feel amused (D’Arms & Jacobson, 2000a). I distinguish “not funny” as indicating *a negative emotional state with high arousal* from “not funny” as indicating *a neutral state*. I point out that when people have an intramural reason not to feel amused by perceiving the moral features of the joke (as opposed to when they consider the consequences of their reaction to the joke), they are in a negative emotional state with

it. It is a happy coincidence that this legitimate reason happens to be a kind of moral reason (and this is why in our arguments later I can make morality relevant in determining the funniness). To see the distinction, imagine a possible world where creatures have the same notion of humor as we do, but instead of morality, they only have “worality”. Whatever that means, they can have a “woral” reason to judge and act. Suppose that an offensive joke can somehow elicit their “woral” reaction in the way that prohibits them from having a reason to feel amused, then this “woral” reason is a legitimate reason from the perspective of humor that has nothing to do with morality.

²² For the sake of convenience, I will speak as if such a group of people indeed exists. The argument can be easily applied to the case where they do not exist. In that case, the claim becomes that even if nobody actually takes moral factors into consideration when judging the funniness of a joke, they are allowed to if they want.

high arousal such as moral disgust or anger.²³ Such a negative emotional state is psychologically incompatible with amusement. According to reason internalism and the principle of “ought implies can,” if one *cannot* feel amused, one *does not have a reason* to feel that way. Thus, one’s intramural reason not to feel amused implies that one does not have a reason to feel amused. The fact that one does not have a reason to feel amused legitimately allows one to judge the joke not funny. By this maneuver, even though moral considerations play a role in the evaluation of funniness, this indirect role does not make one commit the moralistic fallacy.

In section 2 I introduce the charge of moralistic fallacy and a response to it from Jordan and Patridge (2012). In section 3 I provide the main argument, which significantly complements that response. Through sections 4 to 8, I argue for each of the premise in the argument. Alongside the argument, I also make some comments on issues about the relativity and normativity of the concept of humor. I conclude in section 9 that people who have an intramural moral reason against an immoral joke can legitimately claim that the joke is not funny based on its moral offensiveness without committing any fallacy.

3.2. The “Moralistic Fallacy” and a Response

According to the most influential version of the Neo-sentimentalist view on amusement, a joke is funny if and only if one’s feeling amused by it is fitting, where “fitting” means

²³ This is a descriptive claim that I do not support independently. My point is that it is possible that for some people, this is the kind of psychological states they are in. They can be defended using the arguments presented here.

one has a reason to feel that way.²⁴ Accordingly, a joke is not funny if and only if *one does not have a reason* to feel amused by it. In a series of papers, D’Arms and Jacobson argue that moral considerations about the appropriateness of being amused by a joke are irrelevant to whether amusement is fitting (2000a, 2000b, 2014). For example, suppose you find it morally wrong to be amused by a racist joke. D’Arms and Jacobson argue that it would be mistaken to judge the joke not funny based on this moral consideration. They call the inference from moral inappropriateness of amusement to the evaluation of unfunniness a “moralistic fallacy,” which can be summarized as follows:

1. One has a moral reason not to feel amused by an offensive joke.
2. A joke is not funny if and only if one does not have a reason to feel amused by it.
3. Therefore, the offensive joke is not funny.

By giving the label of “fallacy”, D’Arms and Jacobson claim this argument invalid. In my reconstruction of their argument, the fact that *one does not have a reason to feel amused*, as referred to in premise (2), is different from the fact that *one has a reason not to feel amused*, as referred to in premise (1). Therefore, premise (1) does not address the right part of the biconditional in premise (2). Accordingly, the conclusion does not follow from the two premises.

²⁴ I said earlier that fittingness can be interpreted in different ways, as value, correctness, reason, or sui generis. But in all these accounts, there is the consensus that what counts as fitting also counts as good, correct, a reason. Since the discussion around the moralistic fallacy is proceeded along the term of reason, I will also use this term here. But it should be noted that the language chosen is just a matter of being consistent with the literature. All the discussion in terms of reason here can be transformed in terms of fittingness.

I do not dispute the Neo-sentimentalist claim expressed in premise (2).²⁵

However, I think more needs to be said on the distinction between (a) *one has a moral reason not to feel amused* and (b) *one does not have a reason to feel amused*. I argue that in some cases (a) implies (b), and thus in those cases the moralistic fallacy can be avoided.

I would like to introduce this point by discussing the most detailed account in the literature that explains how moral reasons can impede reasons for amusement. Jordan and Patridge (2012) distinguish two mechanisms. *The competing model* is that a moral reason and a funniness reason compete. If the subject finally judges the joke not funny, it is because the moral reason wins against the funniness reason in the competition.

Differently, *the silencing model* is that there is no funniness reason that can compete with the moral reason at all, because the subject's comprehension of the joke silences any could-be reasons for funniness.²⁶

Jordan and Patridge use the following examples to illustrate the distinction. In Scenario 1, you've made an appointment with your friend to see a movie, but on the way

²⁵ I don't claim this is the only plausible view on amusement. My point is, even if this is the right view to adopt, my proposal can still reply to the moralistic fallacy challenge without rejecting this view.

²⁶ The term "silencing" has also been used in works Jordan and Patridge did not mention. For example, Percival (2005) discusses it. In that paper, Percival traces the idea to McDowell (1998). In the literature, the idea about silencing is mostly discussed in virtue ethics. For example, see McDowell & McFetridge, 1978, 26-27. Also, Doris points out that "considerations favoring behavior contrary to virtue are 'silenced' in the virtuous person; although she may experience inducements to vice, she will not count them as reasons for action" (Doris, 2002, 17). Doris continues, "Compare Williams's (1973, 92-3) claim that a person's moral outlook can render some actions 'unthinkable', and Hollis's (1995, 172) gloss of Williams as holding that character sets 'boundary conditions' on the realm of behavioral options. 'Silencing' might be taken to imply that the good person simply does not experience temptations to inappropriate behavior" (Doris, 2002, 176, n.14).

your brother calls to ask for help with his broken car. In Scenario 2, you've made an appointment with your friend to see a movie, but on the way you were hit by a car and had to go to the hospital. They argue that in scenario 1, you have two reasons competing with each other: the reason to go to the movie provided by the appointment, and the reason not to go provided by the fact that your brother is in trouble. In scenario 2, you do not have two reasons competing with each other. The only reason is the reason not to go to the movie: you are hit and need to receive treatment. It would be absurd to say you still have a reason to go to the movie. They add one more thought to make this point more convincing. Suppose the friend with whom you had appointment with went to hospital to see you after she learned about your injury and said, "Why didn't you go to the movie? We had made the appointment!" We would find this remark surprising. How could she even mention the appointment? It is not even a reason under the circumstance! Using this analogy, Jordan and Patridge argue that the reason not to feel amused because the joke is very offensive is like the reason not go to the movie because you were hit. Since it is absurd to say you still have a reason to go to the movie, it is equally absurd to say you still have a reason to feel amused.

This distinction of two models is inspiring. It shows a way to avoid the moralistic fallacy. However, Jordan and Patridge's account does not answer two important questions. First, why do moral considerations sometimes silence and sometimes compete with other considerations? Second, how does silencing happen? In the following, I will answer these two questions by making two further distinctions.

3.3. Avoiding the Fallacy

It is beneficial to display my main argument before explaining it in more detail. It is as follows:

1. People can have an intramural moral reason not to feel amused by immoral jokes.
2. When one has an intramural moral reason not to feel amused, one is in a negative emotional state with high arousal, i.e., moral disgust or anger.
3. The negative emotional state with high arousal (moral disgust or anger) is psychologically incompatible with amusement.
4. Therefore, people who have intramural moral reasons cannot feel amused by immoral jokes. (From 1, 2, and 3)
5. If one cannot feel amusement, one does not have a reason to feel amused.
6. Therefore, people who have intramural moral reasons do not have a reason to feel amused by immoral jokes. (From 4 and 5)
7. A joke is not funny if and only if one does not have a reason to feel amused by it.
8. Therefore, for people who have intramural moral reasons, immoral jokes are not funny. (From 6 and 7)

This argument is valid. If it is also sound, I will have shown that moral reasons can factor into the evaluation of funniness non-fallaciously. In other words, by this argument, people with intramural moral reasons can legitimately claim that immoral

jokes are not funny *because* they are immoral. I will argue for its soundness by explaining each premise in turn.

3.4. Moral Reason: Intramural vs. Extramural

Moral considerations may negatively affect the funniness of a joke in two ways. We may think about the bad moral consequences implied in one's reaction to the joke. The bad moral consequences may include that it would insult or embarrass another member of the audience, make them feel afraid, ashamed or abandoned, and lead to creating, confirming or disseminating discrimination and stereotype. In this case, we have an *extramural moral reason* not to feel amused. We may also think about the negative moral features themselves in the joke without thinking about the consequences of our possible reaction to it. In this case, we have an *intramural moral reason* not to feel amused. See the following joke often cited in the field:

How did a passerby stop a group of black men from committing gang rape? He threw them a basketball.

This is an extremely offensive and thus immoral joke. If one judged this joke not to be funny, we can imagine the following two mechanisms through which one came to this judgment. First, at hearing it, you find there is some wit expressed in this joke that makes it quite clever and funny. However, before you laugh, you give a quick second thought and realize that your laughter might insult and undermine the black people targeted in the joke. With this thought, you then sincerely find the joke not funny anymore. Your judgment is a result of your concern with the potential consequences of your response, of your reflection on what moral stance to take, of showing your clear

moral conscience, and of worrying that you may contribute to spreading certain stereotypes about black people. In this case, you have an extramural moral reason not to be amused because you think about the moral consequences of your possible reaction to the joke. This moral reason outweighs the reason to feel amused, even though without thinking about all those moral consequences the joke appears quite funny to you.

There is a second mechanism through which one can come to the judgment that the joke is not funny. At hearing it, you find this joke outrageous, appalling, disgusting. You don't even get to the point to think about the consequences of your possible reaction. Suppose that at t_1 you understand the joke and understand the moral feature of it and at t_2 you realize the moral consequences of feeling amused. Even if t_1 and t_2 can be very close to each other, you are blocked at t_1 . You think it is not funny because you are angry, disgusted, or shocked. You think to yourself, how could someone come up with a joke like this and intend to create laughter from it? Although you understand the point of the joke, you refuse to judge it funny. You do that not because you want to display your moral acumen, nor because you are concerned about the bad outcome if you express your attitude, nor because you worry that you may contribute to spreading certain stereotypes, but because you just *lack* the inclination to be amused. You are "blown away" by the moral offensiveness of it. You are not influenced by any peer or social pressure, or any force. You are not struggling between two forces. You are not weighing the two reasons pointing in different directions. You find it easy to make a judgment. This moral reason, the reason you have not by considering the moral consequences of

your attitude but by considering the moral features of the joke itself, is what I call the intramural moral reason.

The possibility that sometimes people can have an intramural moral reason while other times only an extramural moral reason is made plausible by the discussion on the sense of humor in the literature. Some scholars argue that the humor of a joke is conditional because jokes require certain background knowledge (Cohen, 1999; de Sousa, 1987b; Woodcock, 2015). This knowledge differs in people with different religious beliefs, cultural and language backgrounds, educational and professional experiences, etc. For example, a person unaware of the stereotype that black men like to play basketball would not find the joke funny; after all, this person just does not understand the joke. I argue that background knowledge can make the humor of a joke conditional in another sense, i.e., these factors may affect one's moral sensibility. Different moral sensibilities in turn may lead to the fact that some people can have an extramural moral reason against the funniness of a racist joke while some other people can have an intramural moral reason.

My proposal of extramural-versus-intramural reasons also dovetails with Haidt's (2001) Social Intuitionist Model and more generally, Moral Foundation Theory (MFT, Graham, et al., 2011; Graham, et al., 2013; Joseph & Haidt, 2007; Haidt & Joseph, 2004). MFT proposes that moral evaluations are based upon several innate, cross-culturally valid psychological systems, i.e. moral foundations. Examples of these moral foundations are care/harm, fairness/cheating, loyalty/betrayal, authority/subversion and sanctity/degradation. Each of the moral foundations is "wired-in" to cope with specific

social problems in the process of human evolution. MFT further proposes that a perceived breach of a moral foundation triggers an immediate, negative emotional response, most commonly in the form of moral disgust or anger (Graham, et al., 2013). More importantly, MFT researchers believe these responses are produced mostly out of an automatic effortless process (the Intuitive System), as opposed to more effortful, deliberative, and cost-benefit based reasoning.²⁷ We do not need to accept the idea that every person will have the same intuitive reaction in the same sort of situation; after all, different cultures may weigh the importance of these moral foundations in different ways. For my purposes, the existence of such an intuitive, emotionally laden process is sufficient. It opens the possibility that people can evaluate an offensive joke without deliberately referring to its moral consequences, i.e. guided by extramural reasons. When people operate under the Intuitive System, they could generate a single, unitary negative emotional response (e.g. moral disgust or anger)—in my terminology, they are guided by intramural moral reasons.²⁸

I would like to make two clarifications regarding this distinction. First, the distinction between the intramural and the extramural moral reasons is not a matter of degree of immorality. In other words, I do not contend that people with an intramural moral reason judge the joke *more* morally outrageous than people with an extramural moral reason do. Rather, in distinguishing the two kinds of moral reasons, I argue that

²⁷ See Haidt (2001) for more discussion.

²⁸ This paragraph was drafted by Kaiyuan Chen.

people attend to *different aspects* of the moral features rather than comparing the immorality in a single scale.

Second, it might be objected that in the literature, scholars have argued for a distinction between relevant and irrelevant reasons for amusement and that no one takes the reason regarding the consequences of one's reaction to be relevant. This objection can be raised from a different perspective: why are extramural moral reasons, which are about the consequence of one's reaction, *moral* reasons? To reply, extramural moral reasons are moral because they concern the *moral* consequences of one's reaction rather than other types of consequences. The typical type of this other kind may be prudential, aesthetic, etc. For example, publicly laughing at an offensive joke may cost one to lose her job. The consideration about this consequence of one's reaction is not moral. The extramural moral reasons are not about such consequences. Rather, they concern whether one's reaction is harmful, disrespectful or vicious in the moral sense. Thus, I agree that considerations about the consequence of one's reaction (whether moral, prudential, or aesthetic) are irrelevant to amusement. What I would like to reject is the claim that *all* moral reasons are irrelevant.

To summarize the distinction, the intramural moral reasons are what one has considering the moral features of the jokes themselves; in contrast, the extramural moral reasons are what one has considering the moral consequences of one's possible reaction to the joke. If there is an extramural moral reason to evaluate a joke, people are liable to being pulled towards different directions by competing reasons. They weigh the value of the extramural moral reason and the value of the witty reason. They quickly do the

calculation. They answer, “It is funny” or “It is not funny,” based on which reason is weightier. In contrast, if there is an intramural moral reason, it blocks or silences the witty reason that could have contributed to the judgment of funniness.²⁹ Thus, in this second case, people are not struggling with different kinds of reasons. Either intramural or extramural moral reasons can make a subject judge that a joke is not funny. But the distinction is important because, as will be shown in the following sections, people with an intramural reason do not commit the moralistic fallacy whereas people who only have an extramural reason do commit that fallacy.

3.5. Negative Emotions with High Arousal: Moral Disgust and Anger

In my main argument in section 2, premise (2) says, “When one has an intramural moral reason not to feel amused, one is in a negative emotional state with high arousal, i.e., moral disgust or anger.” To explain, let me introduce a distinction neglected by D’Arms and Jacobson. According to one shared understanding of emotions, each emotion has at least two features: arousal and valence (Barrett & Russell, 1998; de Sousa, 1987a; Ekman, 1999; Feldman, 1995; Frijda; 1986; Lazarus, 1991; Prinz, 2004; Roseman, 2013; Russell, 2003; Schachter & Singer, 1962; Scherer, 2009; Solomon 1984; Tomkin 1962).³⁰ Arousal concerns physiological and psychological activation. Valence concerns

²⁹ A helpful parallel to the distinction between the extramural and intramural moral reasons is the rebutting defeaters/undercutting defeaters in epistemology. According to Pollock, in evaluating an inference, rebutting defeaters give one a reason to believe the conclusion to be false; undercutting defeaters give one a reason to believe the relationship between the premises and the conclusion is problematic (Pollock, 1967). Parallel to this, we can understand that extramural reasons rebut other reasons for judging the joke funny, while intramural reasons undercut other reasons.

³⁰ The debate about whether emotions are discrete or dimensional is still ongoing in psychology (for example, see Cowen & Keltner 2017, Barrett et al 2018). But this is a debate on how to best characterize those features such as valence and arousal (and maybe others) in emotions. Thus, it will

evaluative judgment.³¹ For example, both anger and annoyance have negative valence, but anger indicates higher arousal than annoyance. Both anger and excitement indicate high arousal, but anger has negative valence and excitement has positive. Amusement is typically characterized as positively valenced with a moderate level of arousal.

Thus, “not feeling amused” may refer to one of two different states. The first is a *neutral state*, where one does not feel anything emotionally. The second is a *negative emotional state with high arousal*, where one feels the opposite of amusement.

Sometimes we use “not amused”/ “not funny” to state a factual matter where amusement/funniness is lacking. In such cases we use the phrases in the emotionally neutral sense. Other times we use the phrases to protest. Imagine a case where you are laughed at by malicious people and you shout at them, “It’s not funny!” In such cases we use the phrases to indicate that we are in a negative emotional state with high arousal. This state of not feeling amused can be labelled as disgust, anger, indignation, outrage, feeling appalled, feeling shocked, etc.³² I propose that the intramural moral reason is paired with a negative emotional state with high arousal and the extramural moral reason

not affect the point here, which just applies the idea that valence and arousal can characterize an emotion.

³¹ The valence is not defined in term of people’s reaction. Some people may have a positive attitude towards a negatively valenced emotion. For example, one might enjoy the feeling of sadness. I do not get into this issue but will stick to the “objective” scale of valence. A helpful analogy is temperature. Temperature can be above or below zero. But people may enjoy only temperatures above zero, only temperatures below, both, or neither.

³² It is worth noting that the feature of negative emotional state with high arousal does not uniquely picks up moral disgust or moral anger. There are other types of emotional states that involve negative evaluation and high arousal, for example, terror, and contempt (in certain occasions). However, these cases are excluded here because it seems implausible to interpret the judgement of “not funny” as expressing these emotions. Thus, even though contempt may be psychologically compatible with amusement, the emotion behind the protest “It’s not funny!” is not.

is paired with the neutral state. This is because it is conceivable that when judging a joke as not being funny, some people feel morally disgusted or angry while other people do not feel anything emotionally. The proposal about the two kinds of reasons can nicely capture this distinction.

3.6. Incompatibility of Emotions

Premise (3) says, “The negative emotional state with high arousal such as moral disgust or anger is psychologically incompatible with amusement.” I will argue for this claim from the *theoretical* (both in philosophy and in psychology) and the *empirical* points of view respectively. But before doing that, I want to make a clarification about the specific emotion types I choose to discuss, i.e., moral disgust and anger.

According to my survey of studies on negative reactions towards immoral jokes, the most common emotion type is moral disgust. There have been some complications. First, in psychology, there is debate about whether and to what extent moral disgust is semantically equivalent to moral anger. For example, lay people may use different words to denote the same emotional state (for the semantical equivalence thesis, see Nabi, 2002; Simpson, Carter, Anthony, & Overton, 2006; for the differentiation thesis, see Hutcherson & Gross, 2011; Rozin, Lowery, Imada, & Haidt, 1999; for the middle-ground thesis, see Gutierrez, Giner-Sorolla, & Vasiljevic, 2012). I will not take stance on this issue. Rather, I will argue as if moral disgust is distinct from anger and anger is a possible reaction to an immoral joke. Thus, I will discuss the two emotions separately. This way, even if it turns out that moral disgust and anger are semantically equivalent, my argument is still tenable. Second, mostly in philosophy, it has been pointed out that

the anger towards immoral behaviors should be divided into two categories. If the subject herself or someone close enough to her is the target of the immoral joke, anger is the proper word. If the target of the immoral joke is not self-related, the proper word is indignation (Descartes [1649] 1985, 397-99; Strawson, 1974). Additionally, regarding this self-other division, some psychological works show that anger is elicited in self-related cases and reduces dramatically in cases unrelated to oneself while moral disgust is more consistent across cases (Hutcherson & Gross, 2011). For this reason, I will take moral disgust as the primary emotion for this discussion, even though I will still talk about anger/indignation, given that it is also mentioned in the literature. Third, it is conceivable that other negative emotional states with high arousal might also be present in response to an immoral joke, such as the feeling of shock, distress, or displeasure. Since they are not the main emotional states discussed in the literature, I will only mention them in passing. However, as I will show briefly, similar arguments apply to them as well.

3.6.1. The Theoretical Point of View

From the theoretical point of view, many scholarly works on the nature of amusement, moral disgust, and anger indicate the incompatibility thesis (i.e., amusement is psychologically incompatible with moral disgust or moral anger) is true. Moral disgust is elicited by perceived moral violations or moral offenses. It is a reaction towards objects or events that threaten social order or moral norms (Rozin, Haidt, McCauley, 2008; Rozin, Lowery, Imada & Haidt, 1999; Tybur, Lieberman, Kurzban & DeScioli, 2013). From a functionalist point of view, moral disgust serves the function of signaling that

something is, prima facie, morally wrong and also motivating people to correct the moral wrong.³³ Moral disgust with high arousal involves a painful, distressful feeling that generally motivates people to avoid and withdraw. On the other hand, amusement involves pleasure and in general motivates one to approach rather than withdrawal (Prinz, 2004; Roberts, 2003). Given the theoretical conceptualization of moral disgust and amusement, it is hard to see how one can enjoy something while feeling pain and distress and having the desire to get away from the wrong-doer, in our context, the joke teller.

Let me consider anger and amusement now. Some psychologists have proposed several ways in which anger and moral disgust are distinct. First, according to the CAD triad model (Rozin Lowery, Imada, & Haidt, 1999), anger responds to violations of autonomy (disrespecting personal rights or freedom) while moral disgust responds to violations of divinity (disrespecting the sacredness of God or causing degradation or pollution to oneself or another). Second, anger is elicited when a moral offense is self-relevant and decreases in cases that are less self-relevant, while the reverse is true for moral disgust (Batson, Chao, & Givens, 2009; Hutcherson & Gross, 2011). Third, anger responds to cues that involves intentional harm, while moral disgust responds to violations of norms related to the body and is not sensitive to intentionality (Russell & Giner-Sorolla, 2011). Fourth, on the dimension of attach/approach and avoid/withdraw,

³³ I say “prima facie” because a functionalist can recognize that feeling moral disgust does not necessarily imply moral wrongness. For example, a person’s feeling disgusted by homosexual people does not imply that homosexuality is actually morally wrong.

anger and moral disgust lie at the opposite ends, i.e., anger motivates approach and disgust motivates withdrawal (Mackie, Devos, & Smith, 2000). Fifth, anger is easier to remedy since it is more concerned with specific actions while moral disgust is not strongly associated with any overt behavior but is predicative of attributing an immoral character to the object of the disgust (Hutcherson & Gross, 2011; Russell & Giner-Sorolla, 2011).

As I will argue now, none of these proposed differences will affect the incompatibility thesis. Aristotle's definition of anger is "[a] desire, accompanied by [mental and physical] distress, for conspicuous retaliation because of a conspicuous slight that was directed, without justification, against oneself or those near to one" (Aristotle, 1991, 124 [1378a]). This concept of anger has been adopted in contemporary theories. Contemporary philosophers of emotions argue that when one is angry, she perceives an offense and takes some party to be culpable or blameworthy (Nussbaum, 2016; Roberts, 2003); she thinks the offending party has frustrated her intentionally (Prinz, 2004); and she regards the offending party as unjustifiably depreciating someone's position (Ben-Ze'ev, 2000). In the psychological study of emotions, functionalists propose that anger indicates that a goal has been undermined and motivates one to remedy the situation (Lench, Tibbett, & Bench, 2016). Thus, even if one's reaction to an immoral joke is anger rather than moral disgust, one would still not feel amusement. After all, it is hard to see how one can enjoy something while feeling pain and distress and desiring to blame the joke-writer and repair the damage created by the joke.

Now consider the proposal that when the party targeted is not related to oneself the proper emotion elicited is indignation. The argument about incompatibility between anger and amusement naturally applies to indignation and amusement. Like anger, a person who feels indignation also feels distress, also perceives an offense, also takes some party to be culpable or blameworthy, also thinks the offending party has frustrated some people intentionally, also regards the offending party as unjustifiably depreciating someone's position. The function of indignation may also be to remedy the damage, although the damage is not related to oneself. Like the case of anger, it seems indignation is psychologically incompatible with amusement.³⁴

3.6.2. The Empirical Point of View

To back these points, now let us look at the incompatibility thesis in the empirical literature. The relevant part of the literature is on mixed emotions, which are defined as “the co-occurrence of positive and negative affects” (Larsen & McGraw, 2014, 263). The core question of this ongoing debate is whether or not certain oppositely valenced emotions are mutually exclusive and thus can be experienced simultaneously. Some theoretical models of emotion, such as Russell's (1980) valence-arousal model, rejects the possibility that emotional states of opposing valence can coexist. While there have also been works on the possible co-occurrence of two seemingly contradictory emotions, this research is usually focused narrowly on certain emotion pairs such as happiness-

³⁴ In all these cases, I do not discuss the degree of the emotions. That is because one aspect that I mentioned, arousal, is roughly what we in ordinary language say about the degree.

sadness (Larsen & McGraw, 2014),³⁵ guilt-pleasure (e.g. the feeling after impulse purchases or watching pornography (Larsen & McGraw, 2014)), fear-pleasure (e.g. the feeling a fan of horror movie has after watching a horror movie (Andrade & Cohen, 2007)), pathogenic disgust-amusement (e.g. the feeling after watching a show where the character eats dog feces (Hemenover & Schimmack, 2007; Larsen & McGraw, 2014))³⁶, tiredness-relaxation (e.g. the feeling a long-distance runner has after completion of a race (Larsen & McGraw, 2014)), and hope-fear (e.g. the anticipatory feeling of going to a restaurant for which one knows there are both good and bad reviews (Bee & Madrigal, 2013)).

I have done a thorough research on the literature on mixed emotions and to my knowledge, the pairs of mixed emotions mentioned above exhaust the combinations that have been studied.³⁷ There is no single study on the possibility of feeling moral disgust and amusement or feeling anger and amusement at the same time. It is possible that

³⁵ In a series of studies, researchers asked people how they feel after watching the movie *Life Is Beautiful* and found that 44% of the participants reported mixed emotions of happiness and sadness (Larsen & McGraw, 2014). Similarly, listeners to songs with conflicting cues also reported such mixed emotions (Hunter, Schellenberg, & Schimmack, 2008). Music with conflicting cues can have, for example, fast tempos (which typically elicit happiness) but minor modes (which typically elicit sadness). Other than art works such as film and music, real life events can also elicit mixed emotions. For example, meaningful endings such as graduation can elicit both happiness and sadness (Larsen, McGraw, & Cacioppo, 2001).

³⁶ Note that the *pathogenic disgust* that has been studied in these works is different from the *moral disgust* at issue here. To briefly explain the difference, scholars have proposed that pathogenic disgust is a reaction towards objects that would potentially cause disease, such as feces, dead body, rotten flesh, mold, etc. Differently, as I have explained, moral disgust is a reaction towards objects or events that threaten social order or moral norms. See Rozin, Haidt, McCauley, 2008; Tybur et al., 2013.

³⁷ See especially Berrios, Totterdell, & Kellett (2015) and Larsen & McGraw (2014).

some future empirical research may show that they are actually compatible. However, the current empirical research supports the incompatibility thesis.

Although there is no direct evidence, there are several pieces of indirect evidence supporting the incompatibility thesis. First, McGraw and Warren (2010) conducted experiments on people's perceptions of immoral jokes and concluded that jokes exhibiting malign violations of moral norms only elicit moral disgust and not amusement. According to McGraw and Warren, a violation is benign when "(a) a salient norm suggests that something is wrong but another salient norm suggests that it is acceptable, (b) one is only weakly committed to the violated norm, or (c) the violation is psychologically distant" (McGraw and Warren, 2010, 1142). They argue that benign violations can elicit both amusement and moral disgust but malign violations tend to elicit strictly negative emotions.³⁸ Similar results can also be found in McGraw et al. (2013) and Kochersberger, Ford, Woodzicka, Romero-Sanchez, & Carretero-Dios (2014). For example, Kochersberger et al. (2014) studied whether the degree of identification with women affects one's perception of sexist jokes. They conclude that for both men and women the more hostile sexist one is, the more one finds sexist jokes amusing. Thus, it is plausible to imagine that a person who holds a strong opinion

³⁸ I address a different question from McGraw and Warren. They test the empirical question of what kind of moral violation can produce humor. My paper answers the normative question of whether it is legitimate for people to find an immoral joke not funny because of moral reasons. My distinction of intramural and extramural moral reasons is partly inspired by their study. However, it is not the case that intramural moral reason corresponds to malign violation and extramural moral reason to benign violation. I argue that one can have both kinds of moral reasons in experiencing a malicious joke, but only the intramural ones are relevant for assessing whether the joke is funny.

against sexism would find sexist jokes not amusing but morally outrageous or disgusting.

The second piece of indirect evidence comes from a recent study on emotion categories. Cowen and Keltner (2017) gathered 2185 short video clips—the widest range studied so far – depicting an exceptionally rich variety of emotional situations and analyzed people’s emotional responses to them. They used a mathematical framework to create a semantic space where the emotional response to each of the video clip is characterized with a unique numeric tag. Among their findings, there is no video clip for which the emotional response is constituted by amusement and moral disgust or by amusement and anger. Importantly, they indeed identified several video clips that can elicit mixed emotions such as sadness and joy. This shows that their study is thorough and fine-grained enough to capture some rare emotional experiences. Thus, the fact that no video clip elicits both amusement and moral disgust or both amusement and anger suggests that it is likely that these emotions are mutually exclusive. Although as many as 2185 video clips may still leave out many other types of emotional experiences in real life, this evidence is good enough for now to support the suggestion that amusement is incompatible with moral disgust and with anger. Of course, the empirical result that no participant feels amusement and moral disgust or amusement and anger at the same time still cannot *prove* the incompatibility thesis. After all, proving that would require testing everyone in every situation, which is itself practically impossible. Since we should make arguments based on the best scientific evidence available (Popper, 1956), I argue that the current empirical research supports the incompatibility thesis more than the opposite.

Thus, I conclude that among those people under the state of moral disgust or anger, which are negative emotions with high arousal, amusement is lacking. There is no room for it. In these cases, when asked, “Do you think the joke is funny or not funny?”, people can legitimately answer “not funny.” When they judge that the joke is not funny due to the intramural moral reason not to feel amused, it is conceivable that, instead of indicating that they are in the neutral state, they mean that they feel not amused as a negative emotional state with high arousal, such as moral disgust or anger. The current empirical research on mixed emotions, the work on malicious jokes only eliciting negative emotions, and the theoretical works on the definition of moral disgust, anger, and amusement by both philosophers and psychologists all indicate that when one feels morally disgusted or angry, one cannot feel amused.

3.7. “Can” and “Is”

Premise (5) says that if one cannot feel amusement, one does not have a reason to feel amused. This is one instantiation of the reason internalism proposed by Williams (1981). Reason internalism is the view that one can have a reason to do something only if that reason is grounded in one’s “prior subjective motivational set”; otherwise, the claim that one has the reason is false. If we understand the notion of reason as indicating some kind of normativity, the idea expressed in premise (5) can also find support in the principle of “ought implies can.” The principle says that one ought to do something only if one has the capability of doing it. In our context, since one is psychologically incapable of feeling amused, we cannot say that one ought to, or has a reason to, feel amused. I will not provide independent argument for reason internalism or the principle of “ought

implies can” but will assume their plausibility, an assumption not unreasonable given their status in the philosophical literature. Adopting these plausible views, having a reason to feel certain way implies that one can psychologically feel that way.

Accordingly, if one cannot feel amused, one does not have a reason to feel amused.

Therefore, when one has an intramural moral reason not to feel amused, which indicates that one is in a negative emotional state with high arousal such as moral disgust or anger, one does not have a reason to feel amused. In these cases, I have shown that it is legitimate to infer *that one does not have a reason to feel amused* from *that one has an intramural moral reason not to feel amused*. This legitimacy is granted by the psychological reality of normal human beings. Combining the Neo-sentimentalist premise (7), some people can legitimately judge that these offensive jokes are not funny *because they are offensive*. This is because amusement and moral disgust or amusement and anger in these contexts are incompatible. Because D’Arms and Jacobson fail to make the distinction between *not amused as a neutral state* and *not amused as a negative emotional state with high arousal*, they incorrectly conclude the argument is invalid.

3.8. Putting Things Together

To summarize, this is how I reply to the moralistic fallacy objection. I made two distinctions, the first is between an *extramural moral reason* and an *intramural moral reason* not to feel amused; the other is between *feeling not amused as a neutral state* and *feeling not amused as a negative emotional state with high arousal such as moral disgust or anger*. When one has an extramural moral reason not to feel amused, she may

not be in a negative emotional state with high arousal. In such a case, her reason not to feel amused does not imply that she does not have a reason to feel amused. In contrast, when one has an intramural moral reason not to feel amused, she does not have a mixed emotion but is feeling not amused as a negative emotional state with high arousal such as moral disgust or anger. In these circumstances, her moral reason not to feel amused implies that she does not have a reason to feel amused.

I do not claim that everyone will read the jokes in the ways as I described. It is possible that some people are quite insensitive to the racial discrimination implied in the basketball joke though they are aware of the relevant stereotypes and thus are amused by it, or that some people may find the joke not funny based on the extramural moral reason. My point is that people who read the joke in the way as I described can be defended against the charge of committing the moralistic fallacy.

Now it should be clear how my account answers the two questions left undeveloped in Jordan and Patridge's account. First, why do moral reasons sometimes silence and sometimes compete with other reasons? My account suggests that when people have intramural reasons, moral considerations silence others; when people have extramural reasons, moral considerations compete with others. Second, how do the two mechanisms operate? My account suggests that intramural reasons make people feel negatively with high arousal but extramural reasons do not. While extramural moral reasons do not preclude the possibility of mixed feelings, the intramural moral reasons do preclude mixed feelings because people having this kind of reasons are in a negative emotional state with high arousal such as moral disgust or anger, which is

psychologically incompatible with amusement.³⁹ Thus, people with intramural moral reasons do not have a reason to feel amused. They are not committing the moralistic fallacy.

Now that I have completed our argument, it is important to emphasize on what I have tried to argue and what I did not aim to argue. I claimed that my thesis is *normative*. By that I mean that people who use an intramural moral reason to judge an immoral joke not funny are *justified* in taking morality into consideration. In other words, in facing the charge of moralistic fallacy, these people can be defended for not committing the fallacy. The defense comes from the fact that these people have an intramural moral reason to feel moral disgust or anger (a reason provided by the way they perceive the joke), which makes them psychologically incapable to and thus lack a reason to feel amused. It is important to note that the charge I discuss and disarm is different from another possible charge. i.e., the charge that people who take the offensive component too seriously lack a sense of humor. This question is also normative, but a different one. It asks, “*Should* these people have an intramural moral reason in reading an immoral joke?”, or “*Should* the offensive component in an immoral joke provide one an intramural moral reason?” The answer to this question lies in the debate about the extent to which morality should affect our sense of humor or to what extent we should

³⁹ My argument about the incompatibility of amusement and anger or moral disgust can also address the alternative explanation that one has a *pro tanto* reason to be amused but an all-things-considered-reason not to be amused. As I have shown above, people who have an intramural moral reason not to be amused are in moral disgust or anger, which psychologically precludes them from feeling the emotion of amusement. Adding the internalist point that one does not have a reason if it is not in one’s subjective motivational set, we can see that these people do not have any reason, whether *pro tanto* or else, to feel amused.

be morally sensitive. I did not and do not aim to address this question. By making the distinction between intramural and extramural moral reasons, I only bring attention to the possibility that the offensive component in the joke may provide them an intramural moral reason not to feel amused. Normatively, we can say that armed with our argument, people with an intramural moral reason are justified in *taking morality into consideration when judging the funniness of a joke*; however, it is still an open question whether they are justified in *cultivating a sensitivity to such moral considerations*. In other words, I leave it open whether they are justified in letting morality affect their sense of humor this much.

3.9. Conclusion

In this paper, I explained why in some cases it is legitimate for people to say that some immoral jokes are not funny because they are immoral. I distinguished two kinds of moral reasons: *intramural moral reasons* and *extramural moral reasons*. I showed in an example that people may have an intramural moral reason to judge the joke is not funny. With this distinction, I defended my view against the charge of committing the moralistic fallacy. I pointed out that weighing the extramural moral reason and the witty reason is fallacious but there is another case where people can legitimately find some immoral jokes not funny, specifically, in reading the joke perceived as severely offensive, they are in the *negative emotional state with high arousal such as moral disgust* or *anger* rather than in a *neutral state*. Thus, what they have is an intramural moral reason not to be amused. In such cases, people's reason not to be amused implies that they do not have a reason to be amused.

I would like to make a final note. My argument is not only a reply to the moralistic fallacy, but also contributes to another current debate on amusement and morality in the literature. Strong Comic Moralism is the view that immoral jokes are not funny (Gaut, 2007); Weak Comic Moralism is the view that immoral joke tokens are not funny (Carroll, 2014; Smuts, 2009, 2010); Comic Immoralism is the view that immoral jokes are funnier (Nannicelli, 2014; Shuster, 2013; Woodcock, 2015); Comic Amoralism is the view that morality is irrelevant to the funniness of a joke (D'Arms & Jacobson, 2000a, 2014); and Comic Pluralism is the view that immorality can both positively and negatively affect the funniness of a joke (Sharadin, 2017). As I have shown, my position is that people have legitimate reasons to find an immoral joke not funny in some cases. This view is different from all of the above in the following aspects.

First, my position is best understood as a reply to Comic Amoralism because I argue that sometimes morality is *relevant* to the funniness of a joke in the sense that people can have legitimate moral reasons to find an immoral joke not funny.

But my position is different from both Strong and Weak Comic Moralism because I believe that people have legitimate moral reasons to find immoral jokes not funny *only under certain circumstances*. In other words, I propose that there are different kinds of moral reasons and only one special kind (the intramural) can legitimately factor into the judgment about funniness. I believe that this more fine-grained distinction can explain two plausible but contrasting intuitions. On the one hand, Comic Moralists use examples such as the basketball joke to pump the intuition that the joke is not funny because of its moral features. On the other hand, objectors stick to the intuition that the

aesthetic features of a joke are independent from its moral feature. With my distinction, we can see that Comic Moralists are correct about offensive jokes like the basketball one because these jokes provide an intramural moral reason. As I have shown, this kind of reason is a legitimate reason for the judgment on the funniness. Meanwhile, Comic Moralists are mistaken in thinking that all moral reasons are relevant: extramural moral reasons are not. We can say that both Strong and Weak Comic Moralism are *Universal Comic Moralism*, no matter whether they are talking about joke type or joke token. In contrast, my claim is *Particular Comic Moralism* because it speaks about a special kind of moral reasons.

Lastly, I do not attempt to make any claim about the symmetrical (possible) group of people defended in Comic Immoralism, who also take morality into consideration in judging the funniness of a joke but find a joke *funnier* because of the moral component. My arguments only discuss cases where moral reasons speak *against* the funniness of a joke and leave open the possibility of moral reasons speaking *for* the funniness.⁴⁰ For example, it may be argued that a racist may have a reason to find a joke

⁴⁰ I think my position and Comic Immoralism can be compatible. Here is a brief explanation. My main argument above is that people may have a special kind of moral reasons against the funniness of a joke, reasons that contribute to the fact that they lack any reason to feel amused because of the incompatibility of amusement and moral disgust or anger. It is possible that some people would not feel negatively but positively towards the immoral component of a joke. Thus, these people may have a reason to find the immoral joke funny. However, even though my position may be compatible with Comic Immoralism, it is *different* from it. As I explained in the section on the incompatibility of certain emotions, that argument supports my claim that for some people, if they feel angry, they do not feel amused. But this argument does not support the comic immoralist claim, for which the emotion of anger is absent.

funny because of the racist component of the joke. But this reason is a morality-related reason *for* the funniness, which falls outside the scope of my discussion.

My view is inspired by contemporary studies on emotion from both philosophy and psychology. I contribute to the literature by offering a fine-grained distinction in the camp of Comic Moralism and depicting a more complicated picture about morality and humor. My view also provides justifications for people who find some immoral jokes not funny to fight against the charge of committing a logical mistake. Finally, my account shows that although for non-moral emotions, fittingness assessment is orthogonal to moral justification, the judgment of whether an emotion (such as amusement) is fitting is constrained by background moral considerations.

CHAPTER IV

APPRECIATION AS AN EPISTEMIC EMOTION: HOW FITTING ANGER IS DIFFERENT FROM INJUSTICE APPRECIATION

4.1. Introduction

On May 25, 2020, George Floyd, an African-American man, was killed by police in Minneapolis. This event triggered widespread and long-lasting protests against police racism, among other things. Suppose that Amy, a doctoral student who comes from South Korea and studies biology at a university in Texas, hears about the news. Immediately, Amy feels anger: it was unacceptable injustice. Amy's anger is fitting because it rightly reflects what is actually true, i.e., the event is indeed unjust.

In a recent paper, Amia Srinivasan argues that fitting anger such as Amy's is intrinsically valuable because through it the subject can *affectively appreciate* the injustice (Srinivasan, 2018).⁴¹ This thesis is proposed against the backdrop of a critique of anger as counterproductive in effect. Srinivasan argues that even if anger is not valuable in the instrumental sense, fitting anger can still have intrinsic value. This move changes the focus of the dialogue from the consequences of anger to the emotion itself and thus constitutes a powerful argument against the criticism of anger, especially in the social-political context. What is more, the argument is significant in a broader sense

⁴¹ Srinivasan's terminology is "aptness" rather than "fittingness". But they mean the same thing. Because "fittingness" is a standard term used by value theorists and philosophers of emotion, I will follow suit.

because it exemplifies how fitting emotions in general can be intrinsically valuable: they are valuable because through them we can affectively appreciate the object of those emotions, for example, irrevocable loss for sadness, achievement for pride, uncertainty for anxiety, etc.

However, without a proper account of “affective appreciation”, it is hard to assess the plausibility of this argument. In this paper, I develop an account of appreciation inspired by Srinivasan’s depiction and then use it to evaluate her thesis about fitting emotion and appreciation.⁴² In my account, there are both a “grasping” and a “feeling” component to appreciation. In experiencing the emotion, we aim to gain better *understanding* of the subject matter in an *affective* way. To put it in a slogan, appreciation is the emotionally charged version of understanding. That is, appreciation is an epistemic emotion in which we form a deep relationship of understanding with the object. Epistemic emotions are affective states related to an epistemic end. I propose that the epistemic end of appreciation is understanding. Understanding is more than propositional knowledge because it requires *cognitive control* of the material while knowledge does not.⁴³ Such a requirement means that one has the capability to see the connection between factors, to tell explanatory stories of why one fact causes another, and to draw correct conclusions in counterfactual cases. To count as an emotion, a state

⁴² “Appreciation” also has two other meanings, one is gratitude (e.g. “She really appreciated the help of her advisors”), the other is seeing the worth or value of something (e.g. “The role of women played in history has not been fully appreciated”). I do not discuss these two meanings.

⁴³ In this paper, all the discussion about knowledge refers to propositional knowledge, unless otherwise noted.

must be intentional, motivational, experiential, and physiological. The feature of understanding reveals that the intentional object of appreciation is plausibility or sense-makingness. Appreciation does not motivate us to investigate further but to maintain the valuable epistemic state of understanding. In appreciation, we experience an “aha” or “eureka” moment. Finally, appreciation is accompanied by certain physiological changes.

Given my account of appreciation, merely having a fitting emotion is not sufficient for appreciation. Accordingly, merely feeling fitting anger is not sufficient for affectively appreciating injustice, contrary to what Srinivasan argues. In many examples, even if one can be fittingly angry, she may still lack a proper appreciation and understanding of the situation. My argument is not to deny that fitting anger can be intrinsically valuable, but is to show that the intrinsic value of fitting anger (and of fitting emotions in general) cannot be explicated by the notion of appreciation. This argument is important because it reveals the limitation of the role of fittingness in emotional assessment as well as highlights the importance of appreciation as a higher goal. Experiencing fitting emotion is only a good starting point. We need to keep engaging with the emotional object in order to achieve the full appreciation of it.

My account of appreciation not only informs us about the evaluation of fitting emotions, but also has two other merits. First, it offers a unified account of how appreciation can occur in many fields, the aesthetic, the moral, and the epistemic. Second, it explicates how aesthetic appreciation is connected to but different from aesthetic emotions.

The plan is as follows. In section 2, I specify the epistemic nature of appreciation. In section 3, I show that appreciation has all the paradigmatic features of an emotion. In section 4, I apply my account of appreciation to fitting emotions. In section 5, I discuss how this account unifies various types of appreciation. In section 6, I situate my account in the literature of aesthetic appreciation. I conclude in section 7.

4.2. Understanding as Grasping

When one appreciates, one understands. Epistemologists articulate the idea of understanding by contrasting it with knowledge. One contrasting point relevant to the current discussion is that understanding is a state in which one “grasps” the object, which involves more cognitive work than knowledge does (Grimm, 2006; 2012; Hills, 2009, 2016; Pritchard, 2016).⁴⁴ Literally speaking, when we grasp an object, we are able to “manipulate or tinker” with it (Grimm, 2011, 89). Analogously, we can grasp a ball, a cup, or a lever. In grasping it, we have it under our control (Hills, 2016). By manipulating it, moving it, turning it around, or putting it to one position or another, we get to know what it is like in various states.

Applying the literal meaning of grasping to the mental realm, we may say that grasping a thing involves the ability to cognitively manipulate and control it. Stephan Grimm calls this a “modal ability.” Allison Hills calls it an “the ability of intellectual know-how” or “cognitive control” (Hills, 2016). Specifically, in Grimm’s words, to be

⁴⁴ According to epistemologists, understanding is different from knowledge also on factivity (Zagzebski, 2001; Grimm, 2011), transparency (Zagzebski, 2001), and being subject to epistemic luck (Kvanvig, 2003). These are all irrelevant here because none of them makes the case that understanding is *more* than knowledge.

able to grasp a system “is to be able to anticipate how changes in one part of the system will lead (or fail to lead) to changes in another part” (Grimm, 2011, 89).⁴⁵ As Duncan Pritchard puts it, by understanding, one “gains a proper explanatory grip on how cause and effect are related” (Pritchard, 2016, 34). In other words, we have an explanatory story about how cause leading to an effect. If one understands why, one has the ability to answer the question, “What if things had been different?” (Grimm, 2006), and the vertical follow-up question, “What does the cause have to do with the effect?” (Lawler, 2019; Skow, 2016). To use an example discussed in Pritchard (2009), a child may *know* via the testimony of a firefighter that the house has burned down because of some faulty wiring. But the child does not *understand* the situation if he cannot see the principles that establish the connection between faulty wiring and fire, or tell the explanatory stories of how the faulty wiring leads to the house burning down, or to draw correct conclusions in counterfactual cases, for example, where the cause of fire was the inappropriate use of an extension cord.⁴⁶

⁴⁵ Grimm also discusses “seeing” involved in “grasping”. For simplicity and consistency with other scholars, I just use “grasping”. This will not affect the main point.

⁴⁶ It is worth emphasizing that there might be different states being compared here. They are: *knowing that*, *knowing why*, *knowing how*, *understanding that*, and *understanding why*. A couple of notes can be made. First, the notion of grasping is primarily discussed in the context of *understanding why*. Second, as for *understanding that*, either that it is more or less synonymous to *knowing that* (Brogaard, 2009; Kvanvig, 2009), or *understanding why* is reducible to it (Kvanvig, 2003, 2009). We can conveniently neglect this phrase, because we will be either talking about *knowing that* or *understanding why*. Third, it seems that *knowing why* is analyzable in terms of a group of *knowing that* claims (Grimm, 2006; Hills, 2016; Pritchard, 2014). This allows us to treat *knowing that* and *knowing why* as the same kind of activity. Since *knowing why* involves several *knowing that* claims and thus seems deeper and more demanding, for our purposes, we can talk only about *knowing why* without losing anything important. Thus, we are left with *knowing why*, *knowing how*, and *understanding why*. The account I want to use here is that *understanding why* is more than *knowing why* and is a kind of intellectual *knowing how* because of the notion of “grasping.”

I contend that when we feel appreciation, the epistemic state that we are in is understanding. Let me use the example of appreciating the injustice of American law enforcement institutions to illustrate this point. When we say that we appreciate why an institution is unjust, we have a *grasp* of this evaluative fact. We not only know this fact by hearing about the news that, for example, George Floyd was forced into unconsciousness and then killed without good reason, but also see that there is a drastic racial discrepancy in prison population that cannot be explained in terms of crime rates, that the police use certain kinds of racial profiling that foster conscious and unconscious bias, and that the court system is designed to discourage disadvantaged groups to file or win a lawsuit, etc. In order to achieve understanding, we also need to be able to explain how these factors contribute to the unjust result, what principles establish the connection between the factors and the result, and whether the institution will still be unjust if some contributors are corrected, etc. In Hills' words, in understanding, we have "cognitive control" of the fact that the institution is unjust. Interpreted this way, *appreciation* of injustice is more than mere *knowledge* of the injustice. In appreciation, we form a deep epistemic relationship of understanding with the object: deeper than knowledge.

Why do I interpret appreciation as involving understanding rather than mere knowledge? There are four motivations. First, this connection between appreciation and understanding is consistent with the language use of many epistemologists working on the topic of understanding. In explaining how understanding is different from knowledge, Jonathan Kvanvig says, "It is the internal seeing or *appreciating* of explanatory relationships in a body of information which is crucial to understanding"

(Kvanvig, 2003, 198). Also, Hills says that “if you understand why X is morally right or wrong, you must have some *appreciation* of the reasons why it is wrong” (Hills, 2009, 101). Besides, in Grimm’s summary, Ernest Sosa holds that firsthand understanding in the case of morality “requires *appreciating* the grounds or reasons for a certain judgment” (Grimm, 2019, 10). Grimm also notes that Wilhelm Dilthey and R. G. Collingwood “have argued that in order to understand people we need to *appreciate*, or in some way sympathetically reconstruct, the reasons that led a person to act in a certain way” (Grimm, 2019, 1).⁴⁷ These philosophers seem to think that there is a very close connection between appreciation and understanding.

The second motivation has to do with discussions on moral deference. In this literature, there is the common objection that moral knowledge obtained through deference is fishy while non-evaluative knowledge obtained through deference is not. For example, we will find it dubious if a person remarks, “It is morally wrong to eat animals. Why? Because that is what my advisor said.” Nevertheless, we will not find it dubious if a person remarks, “It is going to rain tomorrow. Why? Because that is what the weather forecast said.” In articulating the difference, some philosophers appeal to the idea that in moral deference, the agent at best has moral knowledge but not moral understanding (Hopkins, 2007; Hills, 2009, 2011; McGrath, 2011).⁴⁸ Putting this together with the idea of appreciation, if appreciation is interpreted as involving just

⁴⁷ All emphases are mine.

⁴⁸ There are also explanations that do not appeal to the distinction between understanding and knowledge, see Hills (2013) for a review. Since these alternative explanations do not bear to my account, I do not discuss them.

knowledge but not understanding, then we have to say that in moral deference, the agent at best has *moral appreciation* but *not moral understanding*. This is less plausible than the claim that in moral deference, one *lacks* moral appreciation. In other words, if it is mistaken for a person to say “I *understand* the moral wrongness of eating animals. Why? Because that is what my advisor says,” it should be equally mistaken for the person to say, “I *appreciate* the moral wrongness of eating animals. Why? Because that is what my advisor says.”

The third motivation comes from the discussion of aesthetic testimony. There is a debate about whether aesthetic knowledge purely based on testimony is possible. Imagine a person who has never seen a movie yet remarks that it is an excellent piece just based on what a reliable movie critic says. Pessimists of aesthetic testimony think this person lacks aesthetic knowledge about the movie. Optimists disagree by holding that aesthetic beliefs based on such a kind of testimony count as aesthetic knowledge.⁴⁹ In my account, appreciation has a cognitive and an affective aspect. For the sake of argument, suppose that the affective aspect has been accepted by everyone (I will argue for it in the next section). The question here is whether the cognitive aspect is best interpreted as knowledge or as understanding. If it is interpreted as knowledge, then optimists must say that testimony plus some affective experiences caused by testimony count as aesthetic appreciation (because testimony can transmit knowledge), while pessimists must deny this for the same reason that they deny aesthetic knowledge is

⁴⁹ Although the epistemic status of this knowledge may be different. See Robson (2012) for a review of aesthetic testimony.

possible through testimony. I want to be neutral between the pessimists and optimists. Thus, the potential result of divergence on aesthetic appreciation between the two camps is not desirable for me.⁵⁰ Since neither group believe that aesthetic understanding can be conveyed by aesthetic testimony, it is less controversial to interpret appreciation as involving understanding.

The fourth motivation will be clearer when I come to the discussion of Srinivasan's work in section 4. Srinivasan proposes the idea that affective appreciation is more than knowledge. The most crucial task of this paper is to assess whether an account of appreciation can do the work Srinivasan requires it to do. Thus, having an account of appreciation that captures as many features as she proposes is desirable. If my account did not include that appreciation is more than knowledge, then even if it is a plausible account by itself, its application to Srinivasan's work would be powerless. In that case, it could be easily objected that although we use the same term "appreciation", the phenomenon she talks about is different from what I am explaining.

⁵⁰ It might seem peculiar how a combination of aesthetic testimony and an affective experience is possible. One may ask: where does the affective experience come from, given that there is only testimony rather than direct acquaintance with the artwork? The point is well taken. I agree that in real-life scenarios, the combination of testimony and a proper affective experience is rare. But my point here is to explore the logical space rather than to argue for real cases. I suppose that the following scenario is logically possible. I haven't seen the movie, you, as a reliable critic, told me that it is a very good movie. Besides, as an art critic, you are not only reliable but also extremely eloquent. You describe meticulously and vividly how exactly you come to the judgment and what specific feelings you have in each part. After hearing your testimony, I not only have the aesthetic belief, but also have the exact proper affective experience. Now, I have both the aesthetic belief and the required affective experience through your testimony. In this hypothetical case, an optimist would have to say that I have an aesthetic appreciation of the movie while a pessimist would have to deny it.

Given these four considerations, I argue that my account of appreciation as involving understanding rather than knowledge is well motivated.

4.3. Appreciation as an Epistemic Emotion

As shown above, when epistemologists link understanding with appreciation, they do not explicitly emphasize the affective part. I want to make the case that not only does appreciation involve understanding, which has a cognitive aspect, but also that appreciation has an affective aspect.⁵¹ Thus the slogan that appreciation is the emotionally charged version of understanding. Although there are disputes on the nature of emotions among philosophers and affective scientists, most of them agree that an emotion has four paradigmatic features. Emotions are intentional, motivational, experiential, and physiological (Scarantino & de Sousa, 2018).⁵² I will show how appreciation has these paradigmatic features.

Being intentional

Emotions are intentional; they are *about* something (de Sousa, 1987a; Prinz, 2004; Tappolet, 2016; Deonna, 2006; Goldie, 2000; Helm 2001; Johnston, 2001; Mitchell, 2019; Montague, 2009; Roberts, 2003; Zagzebski, 2003; Wringe, 2014). For example, one feels afraid *of* the bear, angry *about* the murder of Floyd, and sad *about* the death of

⁵¹ I remain open to the possibility that epistemologists may conclude that understanding has an affective aspect in itself. However, the literature suggests that they do not usually believe this to be so. Here is why I say this. Laura Callahan recently proposed that moral understanding is different from scientific understanding because the former has an affective and motivational aspect that the latter lacks (Callahan, 2018). This suggests that the standard account of understanding does not include an affective aspect.

⁵² Affective scientists usually add a fifth aspect that emotions are accompanied by facial expressions. The philosophical discussions on emotional expressions are relatively less prominent. Many think this feature is not essential. Thus, I leave out this aspect for now.

her pet. In these examples, the bear, the murder of Floyd, and the death of the pet are called “particular objects” of the emotions (Kenny, 1963). Emotions also have “formal objects,” which are evaluative properties that particular objects share in common. For example, one can fear a bear, a tiger, or a wolf. In all these cases, although the particular objects are different, the formal object is the same, i.e., the *fearsomeness*. Similarly, the intentional object of (moral) anger is *injustice*, and that of sadness is *loss*.

As a subset of emotions, epistemic emotions also have intentional objects. These emotions are those affective states that have to do with enquiry and knowledge gain. Common examples include curiosity, interest, trust, doubt, and wonder (Morton, 2010). For example, when we are curious about what “haiku” means, we take this question as having the property of calling for further investigation. Accordingly, the particular object of curiosity is the meaning of the word “haiku,” and its formal object is the *worthiness of investigation*. We may also be curious about other things. While the particular objects are different, the formal object in all these cases is the same. Similarly, the formal object of interest is *interestingness*, of trust is *trustworthiness*, of doubt is *dubiousness*, of wonder is *questionableness*, etc.

For appreciation, think about the following examples:

- a. “Einstein...had a better intuition for physics than for math, and he did not yet appreciate how integrally the two subjects would be related in the pursuit of new theories.”⁵³

⁵³ Walter Isaacson, *Einstein: His Life and Universe*.

- b. “Dickens enters the social and political war, and the first stroke he deals is not only significant but even startling. Fully to see this we must appreciate the national situation.”⁵⁴
- c. “Those who imagine that far less is required to dismantle mass incarceration and build a new, egalitarian racial consensus reflecting a compassionate rather than punitive impulse toward poor people of color fail to appreciate the distance between Martin Luther King Jr.’s dream and the ongoing racial nightmare for those locked up and locked out of American society.”⁵⁵

Einstein did not yet appreciate the connection between physics and math. We need to appreciate the national situation to see Dickens’ contribution. Some people fail to appreciate the extent of racial injustice. In these examples, the particular objects of appreciation are diverse. When the subjects appreciate these particular objects, they form an understanding of them, have a grasp of the descriptions, evaluations, reasons, causes, explanations of the particular object, and are able to see how things are connected and make sense of them. We can say that the formal object of appreciation in these various cases is *the plausibility of the explanatory story of the particular object, or the sense-makingness of it*. Or put simply, *the worthiness of appreciation*.

Being motivational

Emotions motivate. Scholars have proposed that emotions are adapted reactions to specific goal-related situations, i.e., they are evolved to serve certain ends (Ekman, 1992;

⁵⁴ Gilbert Chesterton, *Appreciations and Criticisms of the Works of Charles Dickens*.

⁵⁵ Michelle Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*.

Lazarus, 1991; Lench et al, 2015; Levenson, 2014; Mauss et al., 2005; Pinker, 1997; Scherer, 1984). For example, fear motivates people to flee, fight back, or freeze, all of which would potentially increase one's chance of living. If we call this kind of emotion a "practical emotion," in that it motivates people to take action, epistemic emotions motivate people to gain knowledge. From the evolutionary perspective, knowledge can be useful in the short or long term. Through curiosity, interest, trust, doubt, or wonder, we are motivated to further enquiry and thus promote gain of knowledge and understanding.

Different from the paradigmatic epistemic emotions that motivate us to do further investigation, appreciation does not motivate *further* investigation but motivates us to *maintain* the valuable epistemic state of understanding, because the end of understanding is *already achieved* in appreciation. When one appreciates something, it feels like that one wants to stay in that state. Analogously, it is like the state we experience when we get to the peak of a mountain that was difficult to climb. After all the hard work of climbing, we eventually see how every point is connected in a plausible way. We want to pause, to stay at the top of the mountain, to immerse ourselves in the moment, to look around and down, to think about every piece in the object, to savor it, to try to make even more sense of it. The object that is worth our appreciation usually attracts our attention, so much so that we are motivated to preserve this state. That is the motivational force of appreciation.

Being experiential

Emotions have distinctive phenomenology. Obviously, emotions such as fear, sadness, anger, or joy have a distinctive phenomenal feel. It is hard to describe what exactly the feelings are, other than to say that everyone who has experienced the emotion definitely knows that it feels a certain way. Sometimes, we know what we feel (as in the cliché that when you fall in love, you will know it). Other times, we may mistake, say, anger for sadness. But it is almost certain that one feels something. The experiential aspect is the most characteristic feature of emotions.

Similarly, we feel something in epistemic emotions. Working in the academia, perhaps the experience of curiosity, interest, trust, wonder, and doubt occurs at a daily base.⁵⁶ The feeling of curiosity and interest is like someone pushing you deep down into a rabbit hole. The feeling of trust is like your heart resting assured regarding the issue. The feeling of wonder and doubt are like you find there is something odd. This is all very metaphorical. But the feeling aspect of epistemic emotions should be familiar, just as in practical emotions.

Appreciation also has an experiential aspect. Kvanvig has pointed out that understanding can produce the “aha” or “eureka” experiences (Kvanvig, 2011). In my account, appreciation involves understanding, so a similar experience is also present in appreciation. Think about your experiences with students in the class of Introduction to Logic. At first, students usually feel desperately confused by the formidable rules of

⁵⁶ Of course, this kind of experience is not restricted to academia.

Modus Ponens, Modus Tollens, Hypothetical Syllogism, Disjunctive Syllogism, etc. Once they have practiced enough, they may tell you how finally “it clicked.” There are also expressions to describe the phenomenon in other languages. For example, in Chinese, there is a phrase called “tí hú guàn dǐng” (醍醐灌顶). “Tí hú” is a liquid dairy product that is good for the skin. This phrase literally means that somebody pours the best thing on you from head to toe. It is a metaphor to indicate that one is suddenly enlightened. These expressions of “aha,” “eureka,” “it clicked,” or “tí hú guàn dǐng” indicate that there indeed is an experiential aspect in understanding, and thus in appreciation.

Being physiological

Finally, emotions are accompanied by certain physiological changes in the body. Since most philosophers usually defer to psychologists on this point or take it for granted from ordinary experience, I will do the same. For example, in discussing aesthetic appreciation, Jesse Prinz comments that “Neuroimaging studies have confirmed that emotions are engaged when people look at art, and the brain areas involved include the anterior cingulate, which is known to be involved in bodily regulation” (Prinz, 2007, 59). We may say similar things about appreciation in general. It is an epistemic emotion that can cause some bodily reactions, perhaps such as making us more awake.⁵⁷

⁵⁷ Granted, what Prinz talks about here seems to take aesthetic appreciation as a single type of emotion. Thus, it is unlikely that he treats, like I do, aesthetic appreciation as the case of the epistemic emotion of appreciation that occurs in the aesthetic realm. So perhaps when he talks about the physiological changes, he does not refer to anything about appreciation but to certain aesthetic emotions. But I think it is at least an open question whether the bodily changes accompanied by appreciation just are or are related to the bodily changes accompanied by aesthetic emotions. The main point is that appreciation in general, not just aesthetic appreciation, seems to be embodied. But as I have said, these issues about physiology are all highly speculative.

To sum up, it is plausible that the uncontroversial aspects of emotions, the intentionality, motivational power, experiential qualities, and physiological manifestations, are all present in appreciation. Thus, whatever theory of emotion we have, appreciation qualifies as an emotion. Because it is related to the pursuit of the epistemic end. i.e., understanding, it counts as an epistemic emotion.

4.4. Assessing Fitting Emotions

With this account of appreciation as an epistemic emotion in which we form a deep epistemic relationship of understanding with the object, now we can assess Srinivasan's claim about fitting emotion and appreciation.

Srinivasan specifically discusses the value of fitting anger. In order to do so, she briefly lays out several necessary conditions for anger to be fitting. First, the object of anger has to be genuinely unjust. For example, if I felt angry towards you for stealing my car but forgot that actually you had borrowed it from me with my permission, then my anger is unfitting because what I was angry about was not unjust. In fitting anger, one has to correctly track an injustice. Second, the object must constitute a reason for the subject to be angry, and the subject must know the reason. For example, if, unbeknownst to me, you stole my car, then even though there is a reason for me to get angry, I do not possess it because I do not know it. Third, the subject must be properly motivated by the reason. For example, suppose that I know you lied to me and I am angry. Suppose also that I am angry not because you lied to me but because everything you did makes me angry. Then, my anger in this case is not fitting (Cf. Srinivasan, 2018, 130).

With this characterization, Srinivasan proposes that in feeling fitting anger we *affectively appreciate* the object, i.e. the injustice. According to this interpretation, the affective appreciation of injustice is analogous to *aesthetic appreciation* and is more than *knowing* that the object is unjust. Just as aesthetic appreciation and knowledge are intrinsically valuable in their own ways, affective appreciation, and thus fitting anger, is also intrinsically valuable. This is an important thesis because it offers a way to argue for the value of fitting emotions in general.⁵⁸ Following the model, we can say that fitting emotions are intrinsically valuable because through them we affectively appreciate the object.

My account of appreciation has the features Srinivasan mentions of being analogous to aesthetic appreciation and more than knowledge. Because appreciation involves understanding and understanding is more than knowledge, appreciation is more than knowledge. For the point about being analogous to aesthetic appreciation, I make a promissory note here. I will show this after the discussion of aesthetic appreciation specifically in section 6. For now, the important point is that my account of appreciation implies that, because fitting anger is different from and does not necessarily involve appreciation, the intrinsic value of fitting anger, if any, cannot reside in appreciation. Before I explain this, a linguistic clarification is necessary.

⁵⁸ Srinivasan does not make this general claim.

4.4.1. *A linguistic clarification*

Note that what I account for is “appreciation” instead of “affective appreciation,” which is the term used by Srinivasan. However, this does not affect my argument. Let me explain.

In general, there are two ways to use a modifier. First, a modifier can indicate a *category*. Think about the following conversation.

A: What’s your dog like?

B: It is a black dog.

In this conversation, when B says she has a black dog, she is distinguishing her dog from dogs of other color categories. In saying that, she assumes that there are dogs of other colors and tries to distinguish her dog from those. Differently, a modifier can also indicate the *nature*. Think about another conversation.

A: Do you like your dog?

B: Absolutely! It is a precious dog!

In this conversation, when B says her dog is precious, she does not assume that there are dogs that are not precious and try to distinguish her dog from those. She is merely emphasizing the precious nature of the dog.

Similarly, the modifiers for appreciation can also be used in these two ways. It seems plausible that phrases like “aesthetic appreciation,” “music appreciation,” “poetry appreciation,” and “moral appreciation” should be read as indicating different *categories*. The modifiers in these phrases refer to the appreciation activity that occurs in the realm of aesthetics, music, nature, and morality, respectively.

Things are different in examples such as “cognitive appreciation”, “rational appreciation,” and “affective appreciation.” These phrases do not seem to refer to the appreciation activity of any particular object in one field. Rather, they indicate an aspect of the activity that is general and applicable to many fields. For example, it is intelligible to say that we cognitively appreciate a scientific theory, an artwork, a political system, or a moral situation. The same goes for rational appreciation and affective appreciation.⁵⁹ Presumably, the terms “affective,” “cognitive,” and “rational” all refer to some form of *faculty* we possess that captures corresponding aspects of the *nature* of the object. This is quite different from how we use the terms of “music,” “poetry,” or “morality.”

My account of appreciation captures this difference in terminology. By nature, appreciation has an affective aspect, because it is an emotion. By nature, appreciation also has an epistemic aspect, because it involves understanding. In terms of category, we can talk about appreciation in aesthetics, morality, science, or others. In all these categories, the nature of appreciation is displayed. When Srinivasan talks about “affective appreciation,” it is plausible to interpret her as not proposing a different *kind* of appreciation, but emphasizing the affective *nature* of appreciation. If this is correct, my account of appreciation as being affective and involving understanding would capture her characterization of affective appreciation.

⁵⁹ This is in accordance with Hills’ account that understanding in morality, aesthetics, and in general shares the same structure.

4.4.2. Fitting anger without appreciation

Now it is time to explain why, given my account of appreciation, Srinivasan's claim that fitting anger is intrinsically valuable because it involves affective appreciation becomes problematic. My account implies that appreciation and anger are two different emotions. The intentional object of anger is the injustice, while the intentional object of appreciation is the plausibility of the explanatory story or the sense-makingness. There are clear cases where fitting anger does involve appreciation of an injustice. For example, the fitting anger felt and expressed by Malcom X does indicate that he really understands the situation, has a good grasping of the condition, and appreciates the unjust structure of the world. However, there are also cases where appreciation is lacking in fitting anger. In other words, not all cases of fitting anger are the same.

Think about Amy in the opening example, the biology doctoral student coming from South Korea. Suppose that she was brought up in an environment where she was never exposed to the racial justice issue. Suppose further that after staying in the United States for 5 years, she has just begun to learn about the issue. So far, all she knows is that the United States has a history of slavery and that African Americans as a group have been discriminated against in more or less overt ways. She has made some friends who care about and understand the issue. They sometimes explain the background information to her. She feels sympathetic to the living conditions of African Americans. Thus, when she hears about the news about George Floyd, she immediately feels anger. What's more, her anger is fitting because the event is indeed terribly unjust and her anger is properly motivated by the reason she possesses. However, she still lacks an

appreciation and an understanding of this unjust event. She has not yet figured out why this event elicited so many protests worldwide. When confronted by people who disagree by saying that Floyd is a criminal who deserves punishment, she does not yet have the capability to explain why this is irrelevant, how this single event is just another case among many that reflects the structural injustice of institutions, and why this is a case not just about police brutality but also police discrimination. It is a plausible case where her anger is absolutely fitting yet she fails to appreciate the injustice. Of course, this is not to say that her fitting anger is thus not valuable. It definitely is, but not in the sense that through it she affectively appreciates the injustice. She cannot do that yet. It can be argued, plausibly I think, that fitting anger is valuable because it is the very first important step towards the deeper epistemic state of understanding. It is perhaps a good starting point for her to learn more about the issue. But that is different from saying that in her anger she already achieves that state.

Similar cases occur in other contexts as well. Actually, when one begins to learn about so many different aspects of social injustice, such as racism, sexism, ageism, xenophobia, ableism, it seems natural to feel sympathetic towards the discriminated people and thus feel fitting anger. However, because it is just the beginning of one's learning, it is hard to say that one already has a good understanding of and can appreciate the situation. To get that stage, one needs to keep learning.

Although not reaching appreciation yet, Amy's anger seems innocent and encouraging. However, there are other cases that are not so innocuous. Think about the Internet, where angry voices abound. How many of these express a kind of fitting anger

wherein the voicer *really understands* the situation? It is not too difficult to find expressions of fitting anger in the comments section below the news article that addresses the unjust murder of George Floyd. But how many of these expressions are written just after a glimpse of the headline? And how many of them are posted just because the headline seems to confirm something one blindly copies from others without bothering to investigate further? Finally, how many of the expressions came out just because the writers are under unconscious peer pressure? In the real world, it can happen that a person sees one instance of the injustice – by reading a post, seeing an advertisement, or watching a video clip – and feels fitting anger immediately, without trying to get more information and a more full-fledged understanding of the event. This anger may well be fitting, since the intentional object is indeed the injustice, and the subject indeed has a reason that motivates the anger properly. But it lacks the sense of appreciation proposed by Srinivasan and discussed extensively in this paper.

Thus, fitting anger is not a *sufficient* condition for moral appreciation of injustice. It is not the case, contra Srinivasan, that to get angry *is* to affectively appreciate the injustice (141), nor *is a means* to affectively appreciating the injustice (132). Rather, to get angry *can be* to affectively appreciate the injustice. Sometimes in fitting anger one has appreciation, and sometimes not.⁶⁰ The implication of the distinction between fitting

⁶⁰ At one point, Srinivasan says she focuses on the case of the victim being angry (130). Thus, the examples I give about the third-party anger do not constitute counterexamples to her. But she also says that “the thought that we can only be aptly angry about things that are sufficiently close to us in space and time, or to which we have some specific personal connection, can shade into a troubling moral parochialism” and discusses some examples about the third-party anger (130-131). So I will treat the first-personal and third-party anger as the same.

anger and appreciation of the injustice is important. It shows that fitting anger is not enough for real understanding, that the role of fitting anger is not ultimate but at best transitional, that we need to follow up on our fitting anger with inquiry, and that we need to be vigilant when we stop at fitting anger.

Similarly, appreciation and other emotions can also come apart. For example, think of a small child's fear of an aggressive dog. It is more natural to say that, in this case, the affective state of fear does not necessarily involve appreciation. The fear is fitting since the intentional object is indeed the fearsomeness and the subject indeed has a reason that motivates the fear properly. But it would be odd to say that in this case the child *appreciates* the fearsomeness of the dog. It is just a quick response, albeit a fitting one. The fearsomeness of the dog may be obvious enough (due to our evolutionary history) that any talk of *appreciation* seems improper (since appreciation requires more effort). But if we follow Srinivasan's model, we have to say that in this case, to feel fitting fear is to "affectively appreciate" the fearsomeness. This is another reason in support of the point that not every case of fitting emotion amounts to appreciation.^{61, 62}

⁶¹ I am not alone in arguing that emotions do not necessarily involve understanding. Michael Brady argues that emotions can facilitate our evaluative understanding but they do not themselves involve understanding (Brady, 2013).

⁶² I admit that it is possible that I am expanding Srinivasan's account beyond her intention. She only talks about the specific emotion of anger and does not explicitly apply this model to other emotions. However, if that is the case, I can modify the dialectic without changing my main arguments. In that case, I would not take the discussion about fear as an objection to Srinivasan's account of affective appreciation, but would take it as an extra argument for the view that fitting emotions and the corresponding appreciation are not necessarily the same thing. This is consistent with how I use her work in the bigger picture. I do not see it as a target of critique, but only as a starting point to build further arguments. In this respect, I owe a great deal to her work.

An objection is, given that appreciation involves understanding, my account implies that it is quite difficult to experience appreciation. To reply, first, this is not necessarily a serious problem. Not everyone feels every kind of emotion. Other than the evolutionarily primitive basic emotions, many others are culturally evolved. Scholars talk about emotion as a capability, meaning that we have certain control to cultivate it. Just as not everybody has the capability to understand everything and not everybody has the capability of aesthetic appreciation, not everybody has the capability to appreciate general objects, including injustice. Second, that appreciation is difficult to achieve can even be a good thing. It can act as an ideal. In the case of anger, understanding social injustice is important but hard. It is a long-term project. Because it is possible that we find ourselves fittingly angry without fully understanding the situation yet, we can be urged to devote more to inquiry. Experiencing fitting anger shows our sensitivity to the situation, which is a good starting point. But to get to a comprehensive appreciation, we must go beyond the starting point.

All in all, my account does not go against Srinivasan's point that fitting anger *can be* intrinsically valuable in the sense of demonstrating our appreciation of the injustice. I fully agree that *sometimes* in fitting anger we appreciate the injustice. In these cases, the intrinsic value of appreciation is like the intrinsic value of understanding.⁶³

⁶³ Understanding is intrinsically valuable because it accurately reflects the way things are (Hills, 2016; Grimm, 2012), because we feel pleasure in it (Hills, 2018), or satisfy our natural desire to make sense of the world (Kvanvig, 2011) because it represents the end or closure of inquiry (Kvanvig, 2011), because it involves a cognitive achievement (Pritchard, 2009, 2010), because it is constitutive of a good human life from the perspective of virtue ethics.

But in other occasions, we do not yet appreciate the injustice even though our anger towards injustice is fitting. That is because appreciation is more demanding than fitting anger. Given this, in discussing the intrinsic value of fitting anger, we need to draw a distinction between *appreciation with fitting anger* from *fitting anger per se*. In general, we need to draw a distinction between *appreciation with fitting X* from *fitting X per se*, where X refers to a specific emotion. This is not to deny that fitting emotions per se (without appreciation) are intrinsically valuable but only to say that if they are, their intrinsic value should be found in places other than appreciation.

There is one last advantage of treating appreciation in my way. In arguing for the thesis that to get angry *is* to affectively appreciate the injustice (141), or that it *is a means* to affectively appreciating the injustice (132), Srinivasan seems to suggest that in fitting anger, the *way* that the subject relates to the object (the injustice) is through affective appreciation.

This is a controversial question that Srinivasan does not expand on. As I have mentioned above, in philosophy of emotion, scholars almost all agree that in an emotion the subject is related to the intentional object in *some way*. However, whether it is a relation of representation, presentation, or indication, or instead some sort of sui generis relation, is highly debated (Colombetti, 2014; D'Arms & Jacobson, 2000a; Goldie, 2000; Hufendiek, 2018; Hutto, 2012; Lazarus, 1991; Montague, 2009; Prinz, 2004; Reisenzein, 2009; Roberts, 2003; Scarantino, 2014; Schargel & Prinz, 2018; Tappolet, 2016). When Srinivasan says that in fitting anger one *appreciates* the injustice affectively, she seems to suggest that the subject relates to the intentional object (the injustice) in an

“appreciative” way. But it is not clear how much commitment she has to saying so. She uses aesthetic appreciation as an analogy. But the intentionality of aesthetic appreciation is also ambiguous. As will be clear in section 6, aesthetic appreciation involves a holistic engagement. There, the relationship between the subject and the object (the aesthetic property) can also be representational, presentational, or other. Given the complexity of this issue, I will not take a stance. Actually, my proposal that appreciation is an epistemic emotion avoids this issue. Whatever kind of relationship a subject is in with the emotional object, that kind of relationship is applicable to the very emotion of appreciation, just as it is applicable to emotions such as anger, sadness, or happiness. In my account, appreciation is not an *alternative* to representation, presentation, indication, or other. It is just another emotion. In this respect, my account of appreciation can fulfill the desiderata of being affective and being more than knowledge while making less of an ontological commitment.

So far, I have articulated my account of appreciation as an epistemic emotion in which the subject forms a deep relationship of understanding with the object.

Appreciation is an *emotion* because it has all the paradigmatic features of an emotion: being intentional, motivational, experiential, and physiological. It is an *epistemic* emotion because it is an emotion related to the epistemic activity that the subject achieves understanding of the object. I applied this account to assess Srinivasan’s claim and showed that it is not necessarily the case that through fitting anger we affectively appreciate the injustice. The implication to the question of emotional fittingness is that, given that fitting anger (and fitting emotion in general) can still be shallow

epistemically, we should weaken the claims about the value and importance of an emotion's being fitting. The value and importance of an emotion's being fitting, as implied in Srinivasan's argument, does not belong to fittingness but to appreciation. My account of appreciation not only informs us about the evaluation of fitting emotions, but also has two other merits. I will articulate them in the next two sections.

4.5. The Many Kinds of Appreciation

First, my account can conveniently unify the phenomena of appreciation in many fields and offers a coherent and systematic interpretation of them.

4.5.1. Aesthetic, moral, and epistemic appreciation

Among different types of appreciation, aesthetic appreciation is the most familiar type. It occurs with respect to aesthetic objects. For example, in our appreciation of Mozart's music, we are connected with the artwork by understanding why it sounds beautiful and by capturing how it creates the harmony that sounds so good to our ears and hearts, are motivated to stay in the music to get more time with it, and experience the feeling of "aha" or "eureka".⁶⁴

There is also moral appreciation, which is the appreciation of virtues and vices, obligations and duties, or morally valuable/disvaluable actions. In the example of appreciation of the injustice, we understand the harm done to the victims in a way that

⁶⁴ Linda Radzik mentions a case about Mark Twain. Twain wrote that a poem is like a frog – if you dissect it, it dies. The idea is that, at least with some kinds of art, understanding how it is made lessens one's aesthetic experience. This is an interesting case. In reply, I want to note that the frog analogy probably works best for the aesthetic experience of awe or surprise but does not work for some other aesthetic experiences such as the feeling of beauty, gracefulness, or sublimity. For example, in appreciating a poem and enjoys its aesthetic qualities, understanding how the organic and the unique features of the poem make it aesthetically good seems to be required. I am grateful to Linda for raising this point.

allows us to track explanatory connections, are motivated to continue engaging with the subject matter and hope to understand it even more, and experience this feeling that we finally grasp what's going on. The "aha" or "eureka" moment does not necessarily mean that we think the event is good. Obviously, injustice is bad. But whether it is good or bad is irrelevant. When we realize the unjust fact, we can have a feeling of being enlightened, or "woken up."

Similarly, there is also epistemic appreciation, which occurs in the epistemic realm related to propositions, arguments, or theories. I mentioned the examples that Einstein did not yet appreciate the connection between physics and math and that we need to appreciate the national situation to see Dickens' contribution. In one's experience of appreciating these things, she is able to cognitively control the claim, is motivated to think more about it, and experiences the feeling that she really gets it.

It is important to note that the term "epistemic" means different things. In "epistemic emotion," it conveys that this emotion is an *attitude about an epistemic activity*. To recall the distinction made earlier about a modifier indicating the nature versus a category, here it "epistemic" indicates the nature of this emotion. In this respect, it is in contrast with *practical emotions* such as fear. Differently, in "epistemic appreciation," "epistemic" conveys that this specific kind of appreciation has *epistemic entities as its objects*. It indicates the category of appreciation. In this respect, it is in contrast with aesthetic or moral appreciation.

That an epistemic emotion can have epistemic, aesthetic, or moral objects is not peculiar. The same goes with practical emotions. For example, we can have fear of

getting the math question wrong (epistemic), fear of hurting the innocent child (moral), fear of choosing a poor word in writing a poem (aesthetic). Just as different cases of fear share the formal object of fearsomeness, which provides us a reason to give them a common name “fear,” aesthetic, moral, and epistemic appreciation share the formal object of worthiness of appreciation, which provides us a reason to give them a common name “appreciation.” Thus, these different types of appreciation belong to the same phenomenon, share the same structure, and have a unified analysis.

4.5.2. *Appreciation requires emotion as a prerequisite*

In the above explanation about what understanding in appreciation is, I only described what is *required* for one to appreciate. Now I would like to discuss briefly how to *achieve* this state. The purpose of doing this is to articulate the connection between appreciation and the related emotions involved in the experience.

Hills takes understanding to be an intellectual know-how. Thus, it needs *practice*, like the practice of a skill or a virtue (Hills, 2016). We only need to take the rudimentary message from virtue theory to notice that the practice of a virtue requires, among other things, the cultivation of proper emotions. For example, the virtue of courage requires feeling properly afraid and feeling properly calm in appropriate times respectively. Without fitting emotions, virtues cannot be attained.

Following this model, to achieve appreciation implies that we have fitting emotions. The role of emotions in moral epistemology has gained much attention. Philosophers argue that emotions can inform us about moral knowledge and that we cannot really understand moral properties without ever feeling corresponding emotions

(Brady, 2013; Deonna & Teroni, 2012; Goldie, 2003; Prinz, 2007; Tappolet, 2016). For example, we get to know injustice through moral anger, moral goodness through admiration, moral wrongness through guilt, etc. This last point is especially obvious. If you have never felt guilt, you would not really understand what moral wrongness is. A typical example of such a case is psychopaths, who seem to know about moral rules without being able to feel bad when they break the rules. They are usually said to lack real moral understanding.

Another example that specifically addresses the role of emotions in moral understanding is Callahan (2018). As shown in the discussion of moral deference, some philosophers notice that in moral deference the subject lacks moral understanding. Callahan argues that the specific aspects lacking in moral understanding are the appropriate affective and motivational responses. These responses are available only when the subject works through the moral question by herself. Thus, Callahan views emotional responses as necessary for moral understanding.

Morality is one category of evaluation. Philosophers of emotion argue that emotions can inform us about evaluative knowledge in general. That is, we cannot really understand evaluative properties without ever feeling corresponding emotions. For example, we get to know the fearsomeness through fear. If you have never felt fear, you would not really understand what fearsomeness is. Similarly, we get to know loss through sadness, disgustingness through disgust, etc.

These works show that the answer to the question “How is the epistemic end of moral/evaluative appreciation, i.e., moral/evaluative understanding, achieved?” is partly

that the subject has to cultivate corresponding moral/evaluative emotions. Without these emotions, one cannot achieve moral/evaluative understanding. Since appreciation has understanding as its essential feature, moral/evaluative appreciation cannot be achieved without fitting moral/evaluative emotions.

The case of epistemic appreciation seems less obvious. For example, a worry is that it is possible to understand how a machine works and yet feel bored by it. Surely boredom is not required for understanding. But in such cases, can't we say that this person still understands the object without experiencing the emotion of appreciation? I think it is phenomenologically true that there is a more obvious connection between understanding and appreciation in moral and aesthetic cases than in epistemic cases. In moral or aesthetic cases, one cannot understand the object without appreciating it. In epistemic cases, it seems that one can understand the subject without appreciating it. But I suspect that whether this is because the word "understand" can be used in a more flexible way in our language than the word "appreciate". In the current context, I tend to think that it is more appropriate to say that, for example, "I *know* how this copy machine works," than to say "I *understand* how this copy machine works," if one is not specialized in copy machine design. But if one is an engineer who has expert knowledge of copy machines, it seems not implausible to say that "She understands how this copy machine works". This shows that the word "understand" allows more room of interpretation than the word "appreciate". In this sense, the objection can be replied this way. If it is an expert, then it is less obvious that she understands how a machine works

and feels bored by it. If it is not an expert, then it is more plausible to say that she knows, instead of understands, how the machine works.

I am not aware of works that discuss the issue that emotions are necessary for appreciation of epistemic entities such as theories, arguments, or propositions. But I think this case can indeed be made. As I have mentioned, paradigmatic epistemic emotions such as curiosity, interest, trust, doubt, or wonder usually function to promote knowledge and understanding gain. For example, without curiosity, our understanding tends to be shallower (Morton, 2010). Thus, we can argue that a good epistemic appreciation requires some kind of epistemic emotion such as curiosity or interest. Accordingly, epistemic appreciation requires certain epistemic emotions.

Another point in reply to this objection is that we need to distinguish “feeling the emotion” from “reporting the emotion”. To use the paradigmatic emotion of fear, it is not always the case that when one says “I feel afraid,” she actually feels that way. For example, in a conversation of spiders, if someone asks “Are you afraid of spiders?” and she answers “Yes, I am afraid of spiders,” then she is not necessarily experiencing the fear. But this is not an objection to the claim that fear is an emotion and can be felt. Similarly, even if for an expert on machine designs, if she claims that “I understand/appreciate how the machine works” as a response to a question, then she is not necessarily experiencing the emotion. But this is not an objection to the claim that appreciation is an emotion and can be felt.

A related point worth noting is that, in general, although appreciation requires certain corresponding emotion, I am not claiming that they must always be co-present.

They certainly can. For example, it is conceivable that when one really understands the injustice for the first time, she feels both angry and appreciation of the fact. But this does not always have to be the case. She may have experienced anger and achieved understanding at one point and only experiences appreciation without anger after since. My account of appreciation does not rule this case out. What my account does rule out is the case where one never feels anger but feels appreciation of injustice. Along with the philosophers mentioned above, I do not think this is possible. One has to experience certain emotions to achieve the corresponding appreciation. In this sense, the emotions that are required to experience to get appreciation plays the necessary “developmental” role.

To summarize, there can be many types of appreciation, the aesthetic, moral, and epistemic. In all these cases, appreciation requires proper emotions as the prerequisite. Moral appreciation requires moral emotions, epistemic appreciation requires epistemic emotions, and aesthetic appreciation requires aesthetic emotions.⁶⁵ Whatever realm is at issue, appreciation is a single experience in which we understand the object in the affective way. It is worth emphasizing that appreciation is affective not because these emotions are needed for understanding but because it is an epistemic emotion.

One might object that it seems peculiar that appreciation as an emotion requires another emotion to occur (for example, moral appreciation of injustice requires moral anger, aesthetic appreciation of the sublime requires the feeling of sublime, epistemic

⁶⁵ The case of aesthetic appreciation will be explained in section 6.

appreciation of an argument requires curiosity, and so on). My reply is that although this is not the paradigmatic case when we talk about basic emotions, the case is actually not uncommon. Prinz's theory of blended emotions says that contempt is the blend of anger and disgust (Prinz, 2007). In aesthetics, the feeling of sublime is constructed as a blend of fear and attraction (Budd, 2003; Fisher, 1998; Kirwan, 2005). There are also mixed emotions such as the feeling of bittersweetness (the mix of sadness and joy), thrill (the mix of fear and joy), or guilty pleasure (the mix of guilt and pleasure) (Larsen & McGraw, 2014). These are all examples of single emotional experiences that have specific intentional objects. We can treat appreciation the same way.

4.6. Aesthetic Appreciation and Aesthetic Emotions

Because aesthetic appreciation is a topic that philosophers have written about extensively, I need to explain what my account of appreciation looks like in that context. I argue that my account not only is compatible with mainstream account of aesthetic appreciation, but also explicates the connection and difference between aesthetic appreciation and aesthetic emotion. This is the last merit of my account of appreciation.

To put it simply, aesthetic emotions are those we have in response to aesthetic objects. They can include ordinary emotions that we can have on other occasions, such as sadness, fear, or joy, but also a particular type that belongs to aesthetic experience exclusively. Examples include the feeling of sublime, beauty, grace, elegance, balance, or kitsch (Fisher, 1998; Kirwan, 2005). We can see by their names that these emotions capture the corresponding aesthetic properties. These are also what aestheticians mainly focus on.

Alan Goldman offers the following holistic account of aesthetic appreciation. “Appreciation is both *understanding* the works and *recognizing* and *enjoying* their artistic values, their values as artworks” (Goldman, 2006, 92, my emphasis).

Specifically,

Great works of art ... engage us on every mental level simultaneously. In them we perceptually appreciate pure sound or color, perceptually-cognitively and perhaps affectively grasp formal structure, cognitively apprehend thematic or symbolic content and historical import, emotionally react to expression, imaginatively expand upon the material present before us, and perhaps even volitionally share in pursuing the aesthetic goals of the works. (Goldman, 2004, 101)

Similarly, Hills argues that aesthetic appreciation is

a practice of engagement with the artwork, responding in various cognitive and noncognitive ways to it, and making a judgment of its value based on the qualities which make it valuable. (Hills, 2018, 173)

Malcolm Budd argues that

[a]ppreciation of a work is not a matter of knowing what its aesthetic properties are, but of perceiving them as realized in the work. (Budd, 2003, 392)

Paisley Livingston argues that

[i]t is one thing to have good ground to think a work of art graceful, exciting, funny, and such; it is something else to gauge, appraise, or actively appreciate it as being so, thereby experiencing and in effect partaking of the state of affairs in

which its inherent value – that is, its contemplative or aesthetic experiential value – is instantiated. (Livingston, 2003, 277)⁶⁶

These accounts of aesthetic appreciation as engagement of cognitive and noncognitive mental faculties at the perceptual, affective, imaginative, and volitional level shows why the experience of appreciating and understanding of an artwork is so rich.

In my framework, aesthetic appreciation is the having of the epistemic emotion of appreciation in the aesthetic realm. In this state, one understands the sense-making features of the aesthetic properties, i.e., why it is graceful, beautiful, elegant, or kitschy, together with the corresponding motivational, experiential, and physiological aspects of appreciation. To achieve such an understanding, one is required to experience certain aesthetic emotions. For example, to understand why the artwork is graceful, one needs to have the emotion of feeling the object as graceful. We don't have a specific name for this emotion. But my intuition is that the emotion exists. Thus, my account of aesthetic appreciation as an application of appreciation in the aesthetic realm has two layers. First, we have appreciation, an epistemic emotion that involves affectivity and understanding. Second, we have aesthetic emotions, which are necessary for achieving the understanding. In other words, there are two types of emotions involved. Aesthetic emotions have *specific aesthetic properties* such as gracefulness, beautifulness, elegance, or others as their intentional objects. Aesthetic appreciation has *the plausibility or the sense-makingness of the specific aesthetic properties* as its intentional object.

⁶⁶ Similar accounts can also be found in Carroll (2016), Franzén (2018), Gorodeisky & Marcus (2018), Lord (2018), Scruton (1974), Sibley (1974).

Aesthetic appreciation cannot be achieved without aesthetic emotions. These two layers together capture the spirit of the holistic account proposed in the above-mentioned philosophers.

With this discussion about aesthetic appreciation, I can cash the promissory note I made on the analogy between aesthetic appreciation and affective appreciation proposed by Srinivasan. Because aesthetic appreciation and moral appreciation are two types of appreciation, they share many things in common. First, they share all the features of appreciation, including the intentional, motivational, experiential, and physiological. Second, both require other emotions to be possible. It seems intuitive that in both the aesthetic and the moral case, there is a strong experiential feature. In my account, in aesthetic appreciation, this experiential feature partly comes from aesthetic emotions (such as the feeling of gracefulness); and in moral appreciation, this experiential feature partly comes from moral emotions (such as anger).

4.7. Conclusion

In this paper, I gave an account of appreciation that links relevant issues in philosophy of emotion, aesthetics, epistemology, and ethics. I argued that appreciation is an epistemic emotion in which the subject forms a deep relationship of understanding with the object. According to this account, appreciation is affective and involves understanding. There are varieties of appreciation that occurs in the aesthetic, moral, and epistemic realms. With this account, I argued that although it is possible that in experiencing fitting anger we affectively appreciate an injustice, this is not necessarily the case. Appreciation requires fitting emotion but it also requires more than that. My account of appreciation

captures the ordinary usage of the term, unifies the phenomenon in different areas and offers a general explanatory schema of them, sheds light on the evaluation of fitting emotions, and explicates the connection and difference between aesthetic appreciation and aesthetic emotion.

CHAPTER V

CONCLUSIONS

Fittingness has been regarded as an important concept in philosophy of emotion and value theory. We have convincing thought experiments indicating that the fittingness assessment is distinct. Many arguments about the value of emotions presuppose that the emotions at issue are fitting. Emotions can work well with reason only if they are fitting. Finally, the account of value in terms of emotions is based on the idea of fittingness.

In the meantime, there is also clear disagreement on how to best understand the notion of fittingness. As I see it, the development of the discussion has evolved in two different directions. On the one hand, for those who enter the discussion primarily concerned with emotions, they tend to rest satisfied with the idea that an emotion's fittingness condition has to do with the accuracy of the corresponding evaluative property. There are still challenges to this idea, such as the circularity problem, i.e., an emotion is said to be fitting if and only if it accurately captures the corresponding evaluative property, meanwhile, it is alleged that an emotion can do this if and only if it is fitting. While there are discussions of the circularity problem, scholars tend to deny it is a problem rather than to propose solutions to it. On the other hand, for those who enter the discussion primarily concerned with normativity, they tend to care more about whether the concept of fittingness is normatively fundamental rather than about giving conditions for it. The consequence of this discussion is that, even if we eventually settle

the question about fundamental normativity, we may still not know *when* an emotion is fitting.

This dissertation is mostly motivated by dissatisfaction with both of these directions. On the one hand, specifying the fittingness condition in terms of reflecting the correct evaluative property seems too uninformative to be plausible. On the other hand, moving away from the discussion of the fittingness condition to the metaethical question about normativity seems less and less relevant to developing a philosophy of emotions. My purpose is to show that there are still interesting discussions to be had about fittingness that diverge from these two directions.

Let me recapitulate how the main chapters are meant to contribute to a broader examination of fittingness as having three constraints. In chapter 2, I argued that the notion of fittingness should leave room for metaethical disagreements and thus should not assume that the content of fittingness is universally applicable to everybody in the same sort of situation. In chapter 3, I argued that in some so-called non-moral emotions, i.e., amusement, the moral considerations are not necessarily the wrong kind of considerations, because they can play a role in the fittingness assessment of the emotion. Thus, fittingness assessment is constrained by background moral considerations. In chapter 4, I argued that fittingness is limited in value because it does not have the value that has been attributed to it, i.e., the value of affective appreciation of the object. I argued that while appreciation is an epistemic emotional state in which the subject forms a deep relationship of understanding with the object, fittingness does not involve this deep connection.

There are several points that can be developed further beyond this dissertation. First and foremost, I raised three constraints that an account of fittingness should respect. But I did not develop a positive account of fittingness. I argued that the standard model of fittingness captures the intuitive idea that fittingness assessment is related to the intentional object of the emotion. However, what does this mean? Several philosophers have suggested that the fittingness criterion of an emotion has to do with its nature (D'Arms & Jacobson, 2000a; Schroeder, 2010; Sharadin, 2016). Perhaps an account of fittingness can be developed by looking at the important aspect of the nature of emotions, i.e., their functions.

According to many psychologists, emotions are adapted reactions to specific goal-related situations and thus function to resolve goal-related problems (Ekman, 1992; Frijda, 1986; Lerner & Keltner, 2001; Lazarus, 1991; Lench, Flores, & Bench, 2011; Levenson, 2014; Pinker, 1997; Scherer, 1984). To give some examples, anger signals that our goal has been blocked and motivates us to take actions to remove the obstacle; fear signals that what we perceive can pose a danger or threat and motivates us to take action to restore safety; admiration signals that someone has achieved excellence in some area and motivates us to emulate her (Onu, Kessler, & Smith 2016). Based on these functions, it is possible to argue that an emotion is fitting when it functions properly, or when it occurs in line with the purposes it is set in place for. For example, if my fear is about a toy bear, this fear does not give me useful information about the object vis a vis my goal of avoiding danger. Therefore, my fear is not fitting. Similarly, if my admiration is directed towards a person who actually performs less well than I do

in a task, this admiration mistakenly inclines me to emulate the person and thus does not promote my goal of improving myself. Therefore, my admiration is not fitting. Similar analyses may apply to emotions in general.

While this positive account of fittingness looks promising, it requires more examination. We may find that some emotions do not fit the same pattern and may need to be treated differently. For example, while it makes sense that practical emotions function for some purposes, it is not obvious how aesthetic emotions such as the feeling of sublime, beauty, elegance, or kitsch function. If these emotions are real, and no explanation of what goals they serve to promote is forthcoming, the functional account will fail (or, at least, apply only to a subset of emotions).

Second, we must consider how best to understand the idea of a function. There is the sense of being functional with respect to the evolutionary history of human beings; there is also the sense of being functional within a specific cultural, historical, or personal context. A functional account would have to explain whether it is possible for an emotion to be functional in one sense but malfunctional in another sense and if so how should this be handled. For example, Martha Nussbaum argues that anger has served its function in our evolutionary history but is no longer functional in our contemporary society (Nussbaum, 2016). A functional account of fittingness would have to reply to challenges of this kind.

Third, a functional account will have to explicate the connection between the goals that emotions are set place to promote by evolution or by one's cultural upbringing, and the goals that an agent consciously entertains. For example, when one

feels angry, although anger functions to remove the obstacle that stands in the way, the agent may not be consciously aware of this goal. One may even have a goal of not hurting people, including not using anger to intimidate the object. In such a case, it seems that the anger fulfills its function defined in the evolutionary or cultural sense, but fails to fulfill the goal set by the individual. Shall we say that the anger is fitting or not? The functional account will need to reply to this challenge by delineating the relevant and irrelevant goals.

Another potential challenge that a functional account faces is in what sense this functional notion of fittingness is normative. We want to use the concept of fittingness to criticize or praise one's emotional responses. If fittingness is only about function, one may object that neither criticism nor praise are appropriate. For example, we do not criticize the heart for being malfunctioning. Advocates of this functional account would need to explain what the difference, if any, is between the heart case and the emotion case.

Relatedly, this account is specifically for emotional fittingness. But as mentioned in the introduction, the idea of fittingness is not restricted to the realm of emotions. There are also fitting beliefs and fitting actions. While it is plausible to regard emotions as strongly goal-directed, it would be interesting to ask whether the same framework can be applied to beliefs and actions. For example, the fittingness of an attitude is related to the nature of the attitude. If the functional account is correct, then the nature of emotion is goal-promoting and thus functional. But the nature of belief is epistemic and not

necessarily goal-directed. Thus, the functional picture in the area of beliefs needs to be explicated.

It is also necessary to examine whether this positive account can work well with the constraints I post in the three chapters. In Chapter 1, I argued that any account of fittingness should give some explanation to the apparent varied intuitions about emotional fittingness. A possible challenge for a functional account is how to address my critique to the pre-judging universalist assumption. For example, to the example of admiration towards somebody who fulfills the duty of filial piety, the functional account needs to explain how apparently admiration is fitting in the ancient Chinese society and not fitting in the contemporary society. It is possible to argue that while the function of admiration remains the same (i.e., it motivates self-improving), what counts as improvement has changed across time. This may be the solution to the problem that, on the one hand, fittingness judgment has normative force, and on the other hand, fittingness judgment should be context-sensitive.

In Chapter 2, I argued that the notion of fittingness is not to be confused with a thicker notion of fitting appreciation. The functional account seems to be compatible with this point. The account says that an emotion is fitting if it functions well to promote its goal. It is conceivable that one can have a fitting emotion that functions well without understanding the object of the emotion. To use the example of anger again, it is possible that one feels fitting anger towards injustice (anger being fitting because it promotes one's goal of correcting injustice) without fully understanding what is involved in the

injustice. Or even clearer, one can fittingly fear a dog (fear being fitting because it promotes one's goal of avoiding danger) without understanding the danger it poses.

In Chapter 3, I argued that one can defend oneself by legitimately claiming that a joke is not amusing because it is immoral. This shows that moral reasons can be relevant for emotional fittingness even in non-moral emotions such as amusement. To check whether the functional account of fittingness can accommodate this, we need to examine the functions of amusement. According to some scholars, amusement functions to signal the incongruity of things that violate our mental expectations. By exercising the abilities in unusual and extreme ways in a safe setting, we are prepared for future unexpected situations (Griskevicius, Shiota, & Neufeld, 2010; Shiota, Campos, Keltner, & Hertenstein, 2004). The functional account would have to say that one's amusement by a joke is fitting when amusement fulfills such a function. This seems to imply that whether there is a moral consideration or not is irrelevant. Thus, there appears a tension between the functional account and the moral constraint. To reply to this challenge, the functional account will need to explain how the moral consideration undermines the function of amusement.

In addition to the challenges of developing this positive account, there are other questions in the area that deserve further investigation. First, I argued for three constraints for an account of fittingness. But this is by no means exhaustive: there may be more. For example, in chapter 3, on wrong kinds of considerations, I focused on the single emotion of amusement and its relationship with moral disgust and moral anger. There are many other emotions that may bear interesting and complicated relationship

with each other. Future research could investigate these cases. It would be interesting to see, after many such case studies are done, whether there are patterns among the factors that play a role in constituting the right and the wrong kind of considerations.

Relatedly, I used examples of admiration, disgust, regret, and nostalgia to argue that fittingness should leave room for normative disagreement. I also acknowledged that in some cases the judgment of emotional fittingness is highly agreed upon. Further research could study why the judgments about the fittingness of some emotions (such as admiration, disgust, regret, and nostalgia) seem to vary significantly with culture, social status, age, and other contextual features, while other fittingness judgments do not. We may ask whether there is a pattern to how the fittingness notion differs according to each type of emotion.

Third, there can be discussions on the parallels between the notion of emotional fittingness and epistemic justification and between emotional fittingness and knowledge. The comparison to epistemic justification is complex. On the one hand, it is argued that the distinction between epistemic justification and other kinds of justification (e.g., moral or prudential) lies in the fact that epistemic justification is somehow connected to the truth of the belief while other sorts of justification are not (Cohen, 1984; McHugh, 2014). This seems to resemble the distinction between emotional fittingness and other kinds of appropriateness (e.g., moral or prudential) and may reveal that emotional fittingness is somehow connected to the correctness of the emotions while other sorts of appropriateness are not. In this sense, emotional fittingness resembles epistemic justification. On the other hand, there are cases of epistemically justified false belief and

epistemically unjustified true belief, while cases of “fitting incorrect emotion” and “unfitting correct emotion” seem less plausible. This suggests that emotional fittingness is importantly different from epistemic justification. For the comparison between emotional fittingness and knowledge, as my discussion on Srinivasan’s work shows, fitting emotion seems to involve affective appreciation, which involves more than propositional knowledge. Thus, fittingness must imply evaluative knowledge. In this sense, fittingness seems to resemble knowledge more than epistemic justification. Further investigation can explore whether the analogies or the disanalogies between fittingness and these epistemic concepts are more significant.

Finally, I discussed the idea of fittingness varying across groups and cultures, it would be interesting to explore whether it can be expanded to entities other than human beings. For example, animals arguably have emotions (Panksepp, 2004). Would it make sense to talk about fitting and unfitting emotions in animals? Also, in the age of artificial intelligence, we can imagine the development of emotional robots and other forms of emotional AI. Would the idea of fittingness apply to them as well or not? These are more challenging questions that are also worthy of further research.

REFERENCES

- Alexander, M. (2020). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- An, D. & Chen, K. (2021). Jokes can fail to be funny because they are immoral: The incompatibility of emotions, *Philosophical Psychology*, 34(3), 374-396, DOI: [10.1080/09515089.2021.1874331](https://doi.org/10.1080/09515089.2021.1874331)
- Andrade, E. B., & Cohen, J. B. (2007). On the consumption of negative feelings. *Journal of Consumer Research*, 34(3), 283-300.
- Antonaccio, M. (2001). Picturing the soul: Moral psychology and the recovery of the emotions. *Ethical theory and moral practice*, 4(2), 127-141.
- Aristotle (1991). *On Rhetoric*, trans G.A. Kennedy, Oxford, Oxford University Press.
- Bagnoli, C. (Ed.). (2011). *Morality and the Emotions*. Oxford University Press on Demand.
- Barrett, L. F., Khan, Z., Dy, J., & Brooks, D. (2018). Nature of emotion categories: comment on Cowen and Keltner. *Trends in Cognitive Sciences*, 22(2), 97-99.
- Barrett, L.F., & Russell, J.A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74, 967-984.
- Batson, C. D., Chao, M. C., & Givens, J. M. (2009). Pursuing moral outrage: Anger at torture. *Journal of Experimental Social Psychology*, 45(1), 155-160.
- Bee, C. C., & Madrigal, R. (2013). Consumer uncertainty: the influence of anticipatory emotions on ambivalence, attitudes, and intentions. *Journal of Consumer Behaviour*, 12(5), 370-381.

- Bell, M. (2013). *Hard feelings: The moral psychology of contempt*. Oxford University Press.
- Ben-Ze'ev, A. (2000). *The subtlety of emotions*. MIT Press.
- Berrios, R., Totterdell, P., & Kellett, S. (2015). Eliciting mixed emotions: a meta-analysis comparing models, types, and measures. *Frontiers in psychology*, 6, 428.
- Boyd, R. (1988). How to be a Moral Realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Cornell University Press. pp. 181-228.
- Brady, M. S. (2013). *Emotional insight: The epistemic role of emotional experience*. OUP Oxford.
- Brink, D. O. (1984). Moral realism and the sceptical arguments from disagreement and queerness. *Australasian Journal of Philosophy* 62(2), 111 – 125.
- Brogaard, B. (2009). What Mary did yesterday: Reflections on knowledge-wh. *Philosophy and Phenomenological Research*, 78(2), 439–467.
- Budd, M. (2003). The acquaintance principle. *British Journal of Aesthetics* 43(4), 386-392.
- Callahan, L. F. (2018). Moral Testimony: A Re-Conceived Understanding Explanation. *Philosophical Quarterly* 68(272), 437-459.
- Carroll, N. (2014). Ethics and comic amusement. *British Journal of Aesthetics* 54(2), 241-253.
- Carroll, N. (2016). Art Appreciation. *Journal of Aesthetic Education* 50(4), 1-14.

- Chappell, R. (2012). Fittingness: The sole normative primitive. *The Philosophical Quarterly*, 62, 684–704.
- Chesterton, G. K. (1911). *Appreciations and Criticisms of the works of Charles Dickens*.
- Cohen, S. (1984). Justification and truth. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 46(3), 279-295.
- Cohen, T. (1999). *Jokes: Philosophical Thoughts on Joking Matters*. University of Chicago Press.
- Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*, Cambridge, MA: MIT Press.
- Cowen, A.S. & D. Keltner (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients, *Proceedings of the National Academy of Science*, 114, E7900-E7909.
- D’Arms, J. (2013). Value and the regulation of the sentiments. *Philosophical Studies* 163(1), 3-13.
- D’Arms, J., & Jacobson, D. (2000a). The moralistic fallacy: On the “appropriateness” of emotions. *Philosophy and Phenomenological Research* 61(1), 65-90.
- D’Arms, J., & Jacobson, D. (2000b). Sentiment and value. *Ethics*, 110(4), 722-748.
- D’Arms, J., & Jacobson, D. (2003). VIII. The significance of recalcitrant emotion, *Royal Institute of Philosophy Supplement* 52, 127-145.
- D’Arms, J., & Jacobson, D. (2014). Wrong kinds of reason and the opacity of normative force. *Oxford Studies in Metaethics*, 9, 215-242.

- Dancy, J. (2000). Should we pass the buck? In A. O'Hear (Ed.), *The good, the true, and the beautiful*, Cambridge: Cambridge University Press, 159–173.
- de Sousa, R. (1987a). *The Rationality of Emotion*. MIT Press.
- de Sousa, R. (1987b). 'When is it wrong to laugh?' in J. Morreall (ed.) *The Philosophy of Laughter and Humor*, Albany, NY: State University of New York Press, 1987, 226–249.
- Deigh, J. (1994). Cognitivism in the theory of emotions. *Ethics* 104(4), 824-54.
- Deonna, J. A. (2006). Emotion, perception and perspective. *Dialectica* 60(1), 29-46.
- Deonna, J. A., & Teroni, F. (2012). *The emotions: A philosophical introduction*. Routledge.
- Deonna, J. A., & Teroni, F. (2015a). Emotions as Attitudes. *Dialectica* 69(3), 293-311.
- Deonna, J. A., & Teroni, F. (2015b). Value and emotion. *Handbook of Value. Perspectives from economics, neuroscience, philosophy, psychology, and sociology*, 155-173.
- Descartes, R. [1649] 1985. *The Passions of the Soul*. In *The Philosophical Writings of Descartes*. Trans. J. Cottingham, R. Stoothoff and D. Murdoch, I: 325–404. Cambridge: Cambridge University Press.
- Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge University Press.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, 6(3-4), 169-200.
- Ekman, P. (1999). Facial expressions. *Handbook of cognition and emotion*, 16, 301-320.

- Feldman, L.A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, 69, 153–166
- Fisher, P. (1998). *Wonder, the rainbow, and the aesthetics of rare experiences*. Harvard University Press.
- Franzén, N. (2018). Aesthetic Evaluation and First-Hand Experience. *Australasian Journal of Philosophy* 96(4), 669-682.
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Gaut, B. N. (2007). *Art, Emotion and Ethics*. Oxford University Press.
- Gert, J. (2016). A Fitting End to the Wrong Kind of Reason Problem. *Ethics*, 126(4), 1015-1042.
- Gertken, J., & Kiesewetter, B. (2017). The right and the wrong kind of reasons. *Philosophy Compass* 12(5): e12412.
- Goldie, P. (2000). *The emotions*. New York: Oxford University Press.
- Goldie, P. (2004). Emotion, feeling, and knowledge of the world. In Robert C. Solomon (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*. Oxford University Press.
- Goldman, A. H. (2006). The experiential account of aesthetic value. *Journal of Aesthetics and Art Criticism* 64(3), 333–342.
- Gorodeisky, K. & Marcus, E. (2018). Aesthetic Rationality. *Journal of Philosophy* 115(3), 113-140.

- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55-130). Academic Press.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of personality and social psychology*, 101(2), 366.
- Greenspan, P. S. (1988). *Emotions and Reasons: An Enquiry Into Emotional Justification*. Routledge.
- Grimm, S. R. (2006). Is understanding a species of knowledge? *British Journal for the Philosophy of Science*, 57(3), 515–535.
- Grimm, S. R. (2011). Understanding. In D. Pritchard & S. Bernecker (Eds.), *The Routledge companion to epistemology* (pp. 84–94). New York, NY: Routledge.
- Grimm, S. R. (2012). The value of understanding. *Philosophy Compass*, 7(2), 103–117.
- Grimm, S. R. (2019). Varieties of understanding, in Grimm (eds.), *Varieties of Understanding: New Perspectives From Philosophy, Psychology, and Theology*. Oup Usa.
- Griskevicius, V., Shiota, M. N., & Neufeld, S. L. (2010). Influence of different positive emotions on persuasion processing: a functional evolutionary approach. *Emotion*, 10(2), 190.
- Guyer, P. (2008). *Knowledge, Reason, and Taste: Kant's Responses to Hume*, Princeton: Princeton University Press.

- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55-66.
- Harman, G. (1996). Moral relativism and moral objectivity.
- Helm, B. W. (2001). *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge University Press.
- Hemenover, S.H., & Schimmack, U. (2007). That's disgusting! . . . , but very amusing: Mixed feelings of amusement and disgust. *Cognition & Emotion*, 21, 1102–1113.
- Hills, A. (2009). Moral testimony and moral epistemology. *Ethics*, 120(1), 94–127.
- Hills, A. (2011). Moral epistemology, in Michael Brady (ed.), *New Waves in Metaethics*. Palgrave Macmillan.
- Hills, A. (2013). Moral testimony. *Philosophy Compass*, 8(6), 552-559.
- Hills, A. (2016). Understanding why. *Noûs*, 50(4), 661–688.
- Hills, A. (2018). Aesthetic Understanding. *Making sense of the world: New essays on the philosophy of understanding*, 160-176.
- Hollis, M. (1995). The shape of a life. *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge.
- Hooker, B., & Stratton-Lake, P. (2006). Scanlon versus Moore on goodness. In T. Horgan, & M. Timmons (Eds.), *Metaethics after Moore* (pp. 149–168). Oxford: Oxford University Press.

- Hopkins, R. (2007). What is wrong with moral testimony? *Philosophy and Phenomenological Research* 74 (3):611-634.
- Howard, C. (2018). Fittingness. *Philosophy Compass* 13(11): e12542.
- Howard, C. (2019). The Fundamentality of Fit. *Oxford Studies in Metaethics* 14.
- Howard, S. A. (2012). Lyrical Emotions and Sentimentality. *Philosophical Quarterly* 62(248), 546-568.
- Hufendiek, R. (2018). Explaining Embodied Emotions—With and Without Representations, *Philosophical Explorations*, 21(2), 319–331.
- Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition & Emotion*, 22(2), 327-352.
- Hurka, T. (2001). *Virtue, vice, and value*. Oxford: Oxford University Press.
- Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of personality and social psychology*, 100(4), 719.
- Hutto, D. D. (2012). Truly Enactive Emotion, *Emotion Review*, 4(2), 176–181.
- Jacobson, D. (2011). Fitting Attitude Theories of Value. *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/>.
- James, W. (1884). What is an Emotion? *Mind*, 9, 188.
- Johnston, M. (2001). The authority of affect. *Philosophy and Phenomenological Research* 63(1), 181-214.

- Jordan, A. & Patridge, S. (2012). Against the Moralistic Fallacy: A Modest Defense of a Modest Sentimentalism about Humor. *Ethical Theory and Moral Practice* 15(1), 83-94.
- Joseph, C., & Haidt, J. (2007). The moral mind: How 5 sets of innate moral intuitions guide the development of many culture-specific virtues, and perhaps even modules. *Carruthers, P.*
- Joyce, R. (2016). Moral Anti-Realism. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/moral-anti-realism/>.
- Kant, I. (1797/1996). *The Metaphysics of Morals*. Cambridge University Press.
- Kavka, G. S. (1983). The toxin puzzle. *Analysis*, 43(1), 33-36.
- Kenny, A. (1963). *Action, Emotion and Will*. Wiley-Blackwell.
- Kirwan, J. (2014). *Sublimity: The non-rational and the rational in the history of Aesthetics*. Routledge.
- Kochersberger, A. O., Ford, T. E., Woodzicka, J. A., Romero-Sanchez, M., & Carretero-Dios, H. (2014). The role of identification with women as a determinant of amusement with sexist humor. *Humor*, 27(3), 441-460.
- Kurth, C. (2015). Moral Anxiety and Moral Agency. *Oxford Studies in Normative Ethics* 5, 171-195.
- Kurth, C. (2018). *The Anxious Mind: An Investigation into the Varieties and Virtues of Anxiety*. Cambridge, MA: MIT Press.

- Kvanvig, J. (2003). *The value of knowledge and the pursuit of understanding*.
Cambridge, UK: Cambridge University Press.
- Kvanvig, J. (2009). The value of understanding. In Pritchard, Haddock & Millar
(eds.), *Epistemic Value*. Oxford: Oxford University Press. pp. 95--112.
- Kvanvig, J. (2011). II—Jonathan L. Kvanvig: Millar on the Value of
Knowledge. *Aristotelian Society Supplementary Volume* 85(1), 83-99.
- Larsen, J. T., & McGraw, A. P. (2014). The case for mixed emotions. *Social and
Personality Psychology Compass*, 8(6), 263-274.
- Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad
at the same time?. *Journal of personality and social psychology*, 81(4), 684.
- Lau, D. C. (ed.) (2000). *Confucius: The Analects*. Columbia University Press.
- Lawler, I. (2019). Understanding why, knowing why, and cognitive
achievements. *Synthese*, 196(11), 4583-4603.
- Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.
- Lench, H. C., Bench, S. W., Darbor, K. E., & Moore, M. (2015). A functionalist
manifesto: Goal-related emotions from an evolutionary perspective. *Emotion
Review*, 7(1), 90-98.
- Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in
cognition, judgment, experience, behavior, and physiology: a meta-analysis of
experimental emotion elicitation. *Psychological bulletin*, 137(5), 834.

- Lench, H. C., Tibbett, T. P., & Bench, S. W. (2016). Exploring the toolkit of emotion: What do sadness and anger do for us?. *Social and Personality Psychology Compass*, *10*(1), 11-25.
- Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of personality and social psychology*, *81*(1), 146.
- Levenson, R. W. (2014). The autonomic nervous system and emotion. *Emotion Review*, *6*(2), 100-112.
- Livingston, P. (2003). On an apparent truism in aesthetics. *The British Journal of Aesthetics*, *43*(3), 260-278.
- Loeb, D. (1998). Moral realism and the argument from disagreement. *Philosophical Studies*, 281-303.
- Lord, E. (2018). How to learn about aesthetics and morality through acquaintance and deference. *Oxford studies in metaethics*, *13*, 71-97.
- Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: explaining offensive action tendencies in an intergroup context. *Journal of personality and social psychology*, *79*(4), 602.
- Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*(2), 175.
- McDowell, J. H. (1985). Values and secondary qualities. In Ted Honderich (ed.), *Morality and Objectivity*. London: Routledge. pp. 110-129.
- McDowell, J. H. (1998). *Mind, value, and reality*. Harvard University Press.

- McDowell, J., & McFetridge, I. G. (1978). Are moral requirements hypothetical imperatives?. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 52, 13-42.
- McGrath, S. (2011). Skepticism about moral expertise as a puzzle for moral realism, *Journal of Philosophy*, 3. 111–137.
- McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological science*, 21(8), 1141-1149.
- McHugh, C. (2012). The truth norm of belief. *Pacific Philosophical Quarterly*, 93, 8–30.
- McHugh, C. (2014, July). VIII—Fitting Belief. In *Proceedings of the Aristotelian Society*, 114(2pt2), 167-187.
- McHugh, C., & Way, J. (2016). Fittingness First. *Ethics* 126(3), 575-606.
- Mitchell, J. (2019). Emotional Experience and Propositional Content. *Dialectica* 73(4), 535-561.
- Montague, M. (2009). The logic, intentionality, and phenomenology of emotion. *Philosophical Studies* 145(2), 171-192.
- Moore, G. E. (1903). *Principia ethica*. Cambridge: Cambridge University Press.
- Morton, A. (2010). Epistemic Emotions. In Peter Goldie (ed.), *The Oxford Handbook of Philosophy of Emotion*. Oxford University Press. pp. 385--399.
- Na'aman, O. (2020). The Rationality of Emotional Change: Toward a Process View. *Noûs*.
- Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion research. *Cognition & Emotion*, 16(5), 695-703.

- Nannicelli, T. (2014). Moderate Comic Immoralism and the Genetic Approach to the Ethical Criticism of Art. *Journal of Aesthetics and Art Criticism* 72(2), 169-179.
- Nussbaum, M. C. (2001). Upheavals of Thought. *Journal of Religious Ethics* 31(2), 325-341.
- Nussbaum, M. C. (2016). *Anger and forgiveness: resentment, generosity, justice*. Oxford University Press.
- Onu, D., Kessler, T., & Smith, J. R. (2016). Admiration: A conceptual review. *Emotion Review*, 8(3), 218-230.
- Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press.
- Parfit, D. (2011). *On what matters*. Oxford University Press.
- Percival, P. (2005). Comic Normativity and the Ethics of Humour. *The Monist*, 88(1), 93-120.
- Pinker, S. (1997). *How the Mind Works*. New York, NY: Norton.
- Pollock, J. L. (1967). Criteria and our knowledge of the material world, *Philosophical Review*, 76, 28-62.
- Popper, K.R. (1959). *The Logic of Scientific Discovery*, London: Routledge, 2002
- Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. New York: Oxford University Press.
- Prinz, J. (2007). *The emotional construction of morals*. Oxford University Press.
- Pritchard, D. (2009). Knowledge, Understanding and Epistemic Value. *Royal Institute of Philosophy Supplement* 64, 19-43.

- Pritchard, D. (2010). Knowledge and Understanding, in *The Nature and Value of Knowledge: Three Investigations*. Co-authored with Eds. Alan Millar and Adrian Haddock. New York: Oxford University Press, 2010, 3–90.
- Pritchard, D. (2014). Knowledge and understanding. In A. Fairweather (Ed.), *Virtue epistemology naturalized* (pp. 315–327). New York, NY: Synthese Library. 366.
- Pritchard, D. (2016). Seeing it for oneself: Perceptual knowledge, understanding, and intellectual autonomy. *Episteme* 13(1), 29-42.
- Pugmire, D. (1998). *Rediscovering Emotion*.
- Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The strike of the demon: On fitting pro-attitudes and value. *Ethics*, 114(3), 391-423.
- Reisenzein, R. (2009). Emotions as Metarepresentational States of Mind: Naturalizing the Belief-Desire Theory of Emotion, *Cognitive Systems Research*, 10(1), 6–20.
- Roberts, R. C. (2003). *Emotions: An essay in aid of moral psychology*. Cambridge University Press.
- Robson, J. (2012). Aesthetic Testimony. *Philosophy Compass* 7(1),1-10.
- Rønnow-Rasmussen, T. (2011). *Personal value*. Oxford University Press.
- Roseman, I. J. (2013). Appraisal in the emotion system: Coherence in strategies for coping. *Emotion Review*, 5, 141–149.
- Rosen, G. (2015). The alethic conception of moral responsibility. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The nature of moral responsibility: New essays*. Oxford: Oxford University Press, 65-88.
- Rowland, R. (2013). Wrong kind of reasons and consequences. *Utilitas*, 25(3), 405-416.

- Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: a mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of personality and social psychology*, 76(4), 574.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178. doi: 10.1037/h0077714
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145–172.
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, 11, 233-240.
- Scanlon, T. (1998). *What we owe to each other*. Belknap Press.
- Scarantino, A. (2014). The Motivational Theory of Emotions, in Justin D’Arms and Daniel Jacobson (eds.), *Moral Psychology and Human Agency*, Oxford: Oxford University Press, 156–185.
- Scarantino, A., & de Sousa, R. (2018). Emotion. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/emotion/>
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review*, 69(5), 379.

- Schargel, D., & Prinz, J. (2018). An Enactivist Theory of Emotional Content, in Hichem Naar and Fabrice Teroni (eds.), *The Ontology of Emotions*, Cambridge: Cambridge University Press, 110–129.
- Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality and Social Psychology*, 5, 37–63.
- Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23, 1307–1351.
- Schmitter, A.M. (2005). The Passionate Intellect: Reading the (Non-)Opposition of Intellect and Emotion in Descartes, in *Persons and Passions: Essays in Honor of Annette Baier*, ed. J. Whiting, C. Williams, and J. Jenkins, Notre Dame, IN: Notre Dame U. Press, 48-82.
- Schmitter, Amy M. (2016). 17th and 18th Century Theories of Emotions, *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2016/entries/emotions-17th18th/>>.
- Schroeder, M. (2010). Value and the right kind of reason. *Oxford Studies in Metaethics* 5, 25-55.
- Scruton, R. (1974). *Art and Imagination: A Study in the Philosophy of Mind*. St. Augustine's Press.
- Setiya, K. (2017). *Midlife: A Philosophical Guide*. Princeton University Press.
- Shafer-Landau, R. (1994). Ethical Disagreement, Ethical Objectivism and Moral Indeterminacy. *Philosophy and Phenomenological Research* 54, 331–344.
- Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford University Press.

- Sharadin, N. (2016). Reasons Wrong and Right. *Pacific Philosophical Quarterly* 97(3), 371-399.
- Sharadin, N. (2017). In Defense of Comic Pluralism. *Ethical Theory and Moral Practice* 20(2), 375-392.
- Shiota, M. N., Campos, B., Keltner, D., & Hertenstein, M. J. (2004). Positive emotion and the regulation of interpersonal relationships. *The regulation of emotion*, 68.
- Shuster, M. (2013). Humor as an Optics: Bergson and the Ethics of Humor. *Hypatia* 28(3), 618-632.
- Sibley, F. N. (1974). Particularity, Art and Evaluation. *Aristotelian Society Supplementary Volume* 48(1), 1-22.
- Simpson, J., Carter, S., Anthony, S. H., & Overton, P. G. (2006). Is disgust a homogeneous emotion?. *Motivation and emotion*, 30(1), 31-41.
- Sinnott-Armstrong, W. (2002). Moral relativity and intuitionism. *Philosophical Issues*, 12 (*Realism and Relativism*), 305–328.
- Skorupski, J. (2007). Internal reasons and the scope of blame. *Bernard Williams*, 73-103.
- Skow, B. (2016). *Reasons why*. Oxford, UK: Oxford University Press.
- Smuts, A. (2009). Do Moral Flaws Enhance Amusement? *American Philosophical Quarterly* 46(2), 151-163.
- Smuts, A. (2010). The Ethics of Humor: Can Your Sense of Humor be Wrong? *Ethical Theory and Moral Practice* 13(3), 333-347.
- Solomon, Robert 1984. *The Passions: The Myth and Nature of Human Emotions*, New York: Doubleday.

- Srinivasan, Amia (2018). The Aptness of Anger. *Journal of Political Philosophy* 26(2), 123-144.
- Stratton-Lake, P. (2005). How to deal with evil demons: comment on Rabinowicz and Rønnow-Rasmussen. *Ethics*, 115(4), 788-798.
- Strawson, P. F. (1974). *Freedom and resentment, and other essays* (Vol. 595). Egmont Books (UK).
- Sylvan, K. (2012). Truth monism without teleology. *Thought*, 1, 161–163.
- Tappolet, C. (2016). *Emotions, Value, and Agency*. Oxford University Press UK.
- Taylor, G. (1985). Pride, shame, and guilt: Emotions of self-assessment.
- Tomkins, S. S. (1962). *Affect, imagery, and consciousness*. New York, NY: Springer.
- Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological review*, 120(1), 65.
- Walter, I. (2007). *Einstein: His Life and Universe*.
- Way, J. (2012). Transmission and the Wrong Kind of Reason. *Ethics* 122(3), 489-515.
- Way, J. (2013). Value and reasons to favour. In *Oxford studies in metaethics* 8, 27–49.
- Wiggins, D. (1987). A sensible subjectivism? In his *Needs, values, truth*.
- Williams, B. (1973). *Problems of the Self*. Cambridge University Press.
- Williams, B. (1981). *Moral luck: philosophical papers 1973-1980*. Cambridge University Press.
- Wilson, E. E., & Denis, L. (2018). Kant and Hume on Morality. *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.),

URL = <<https://plato.stanford.edu/archives/sum2018/entries/kant-hume-morality/>>.

Wodak, D. (2020). Who's On First? *Oxford Studies in Metaethics* 15.

Wong, D. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press USA.

Wood, A. W. (2008). *Kantian Ethics*, New York: Cambridge University Press.

doi:10.1017/CBO9780511809651

Woodcock, S. (2015). Comic Immoralism and Relatively Funny Jokes. *Journal of Applied Philosophy* 32 (2): 203-216.

Wringe, B. (2014). The Contents of Perception and the Contents of Emotion. *Noûs* 48(1), 275-297.

Yam, K. C., Barnes, C. M., Leavitt, K., Wei, W., Lau, J., & Uhlmann, E. L. (2019). Why so serious? A laboratory and field investigation of the link between morality and humor. *Journal of personality and social psychology*.

Yam, K. C., Christian, M. S., Wei, W., Liao, Z., & Nai, J. (2018). The mixed blessing of leader sense of humor: Examining costs and benefits. *Academy of Management Journal*, 61(1), 348-369.

Zagzebski, L. (2001). Recovering understanding. In S. Matthias (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue* (pp. 235–252). Oxford, UK: Oxford University Press.

Zagzebski, L. (2003). Emotion and moral judgment. *Philosophy and Phenomenological Research* 66(1), 104–124.