

Georgia State University

ScholarWorks @ Georgia State University

Philosophy Theses

Department of Philosophy

8-2023

In Defense of Emotion-Based Approach to Moral Motivation

Yifan Wang

Follow this and additional works at: https://scholarworks.gsu.edu/philosophy_theses

Recommended Citation

Wang, Yifan, "In Defense of Emotion-Based Approach to Moral Motivation." Thesis, Georgia State University, 2023.

doi: <https://doi.org/10.57709/35681246>

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

In Defense of Emotion-Based Approach to Moral Motivation

by

Yifan Wang

Under the Direction of Eddy Nahmias, PhD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2023

ABSTRACT

Some philosophers have advanced emotion-based arguments for motivational internalism by arguing that emotions play constitutive roles in moral judgments and that emotions motivate actions. Jennifer Corns and Robert Cowan object to this approach, arguing that linking emotion and moral motivation does not help advance the debate about motivational internalism, because mental states that are paradigmatically pleasant or unpleasant, including emotions, do not necessarily motivate agents. I argue that their objections are misplaced as emotion-based accounts do not and do not have to rely on the pleasantness or unpleasantness of emotions to make the case that emotions necessarily motivate. I further propose a revised argument for internalism in which I argue that emotions should be understood as action tendencies with functional goals. The revised argument addresses their objections and improves the existing emotion-based arguments, allowing us to explain how emotions motivate agents to perform specific actions prescribed by moral judgments.

INDEX WORDS: Moral motivation, Motivational internalism, Emotion, Action tendency, Goals, Moral judgments

Copyright by
Yifan Wang
2023

In Defense of Emotion-Based Approach to Moral Motivation

by

Yifan Wang

Committee Chair: Eddy Nahmias

Committee: Andrea Scarantino

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2023

DEDICATION

This thesis is dedicated to my grandfather.

ACKNOWLEDGEMENTS

I would like to thank Dr. Eddy Nahmias and Dr. Andrea Scarantino for all their invaluable help and extraordinary patience throughout the process of my writing. I have benefited enormously from their guidance. I especially would like to thank Dr. Nahmias for all the time and effort that he put into advising, reviewing, and commenting on the numerous drafts I wrote, all of which helped me figure out what I really wanted to say.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		V
1 INTRODUCTION		1
2 EMOTION-BASED ACCOUNTS FOR INTERNALISM		3
2.1 Zagzebski’s Pluralist Approach		5
2.2 Prinz’s Sentimentalist Approach		9
3 CORN & COWAN’S OBJECTIONS		14
4 EMOTIONS ARE FUNCTIONALLY MOTIVATING		18
4.1 Response to ‘Liking’ vs ‘Wanting’ Objection		18
4.2 The Missing Piece of the Puzzle		21
4.3 The Revised Argument		22
4.4 P1: Aboutness and Specific Objects		23
4.5 P2: Emotions as Action Tendencies with Goals		27
4.6 P4: From Goals to Specific Actions		31
5 UNWARRANTED GUILT AND MORAL DESPAIR		35
5.1 Unwarranted Guilt		35
5.2 Moral Despair		36
6 CONCLUSION		39
BIBLIOGRAPHY		41

1 INTRODUCTION

After watching a documentary showing the death of numerous marine life due to ocean pollution, especially pollutants made of plastics, Nancy judges that it is wrong for her to carelessly use disposable cups on a daily basis. Consequently, she decides to avoid using disposable cups as much as possible. She then develops a habit of bringing the same reusable cup with her wherever she goes. Both motivational externalists (henceforth, externalists) and motivational internalists (henceforth, internalists) recognize that Nancy's moral judgment that it is wrong to carelessly use disposable cups lead her to switch to reusable cups. But they disagree on whether and how exactly her moral judgment motivates her to do so.

While there are many versions of internalism, internalists generally think that the source of motivation comes from the normative judgment itself. If one genuinely makes a moral judgment that one morally ought to act in certain ways, then that moral judgment by itself is sufficient to motivate one to act accordingly, even if that motivation might be outweighed by other factors. In the case of Nancy, it is her judgment that it is wrong to use disposable cups that motivates her to stop using disposable cups. Externalists, on the other hand, argue that the source of motivation comes from something external to the moral judgment itself (Rosati, 2016). For instance, Nancy might have been raised to have desires to help the environment or to desire to act in a way that will be perceived by others as progressive. It is not the moral judgment itself, but these other desires that motivate her to use reusable cups.

In the internalist camp, philosophers like Linda Zagzebski and Jesse Prinz focus on the connection that moral judgments have with emotions. They argue that emotions are necessary components for (some or all) moral judgments and that emotions motivate us to act. Hence, moral

judgments by themselves necessarily motivate us. Their arguments will be referred to as emotion-based accounts of internalism, hereinafter.

In a recent paper, Jennifer Corns and Robert Cowan (C&C) refer to philosophers like Zagzebski and Prinz as “proponents of ‘the affective appeal’ ” as they appeal to the motivational force of affective mental episodes to defend internalism (71–73).¹ By affect or affective mental episodes, C&C refer to conscious mental episodes that are paradigmatically pleasant or unpleasant, such as emotions, pleasure, pain, and pleasant or unpleasant thoughts (C&C 2021, 73 and 76; Corns 2014, 240). C&C argue that this strategy of appealing to affect to defend internalism ultimately fails. Specifically, C&C think that showing that emotions play constitutive roles in moral judgment is not enough to defend internalism. The reason is that pleasant or unpleasant mental states, including emotions, do not necessarily motivate actions. As they point out, evidence from cognitive science shows that humans can be in states of pleasure or pain without being in a motivational state and vice versa.

I argue that, by reducing the emotion-based accounts to pleasant or unpleasant feelings, C&C have not done emotion-based internalism justice. Neither Zagzebski nor Prinz is saying that it is the pleasure or pain of emotions alone that allows emotions to motivate agents, as they both point to the possibility that emotions are associated with action tendencies that make emotions functionally motivating, an idea that will be further explicated in this thesis.

In Section 2, I will explain Zagzebski and Prinz’s emotion-based accounts. Section 3 will explain the three main objections that C&C raise to emotion-based accounts. In Section 4, I argue that emotion-based accounts do not primarily rely on the pleasant or unpleasant feelings of

¹ In the same paper, C&C also argue against Jonathan Dancy and Antti Kauppinen who think that moral intuitions are affective, which makes intuitions a type of “seeming”. Because Dancy and Kauppinen discuss moral intuitions rather than moral judgments *per se*, in my paper, I will focus on the emotion-based accounts for moral internalism given by Prinz and Zagzebski.

emotions to make the case that emotions necessarily motivate. I further propose a revised argument for internalism where I argue that emotions should be understood as action tendencies with goals. Section 5 will come back to two counterexamples raised by C&C which are meant to show that emotions might not motivate us to act. Finally, I conclude the thesis in Section 6. My arguments will leave emotion-based accounts of motivational internalism in a stronger position, in part by providing a better way to understand how emotions motivate relevant actions.

2 EMOTION-BASED ACCOUNTS FOR INTERNALISM

C&C focuses on the following version of internalism:

“Unconditional Motivational Internalism (UMI): *Necessarily*, if *S* judges that she morally ought to ϕ , then *S* is motivated (at least somewhat) to act in accordance with her judgment” (2021, 72).

The versions of internalism that Zagzebski and Prinz each focus on are similar to UMI, but in a sense, more qualified. At the beginning of her paper, Zagzebski defines internalism (what she calls “motivational judgment internalism”) as the thesis that “all we need do to get [a person] to feel a motive to act on a moral judgment is to get her to make the judgment” (2003, 104-105).² But as we will see later in this section, her actual position is that only some types of moral judgments will motivate the agents—the moral judgments that are expressions of emotions.

Prinz defines internalism as the thesis that “moral judgments *eo ipso* place those who make them in a motivational state to act in accordance with those judgments” (2015, 62). But the nuance here is that emotions are built into his definitions of moral judgments, a point that will also be discussed further in this section. To summarize the positions taken by the emotion-based

² Although Zagzebski uses the term motive, it seems that she uses motives and motivation interchangeably throughout the paper.

internalists, they both think that making a moral judgment or making a certain type of moral judgment is sufficient to motivate a person to act on it. In other words, being motivated to act in accord is the *necessary* consequence of making (some or all) moral judgments.

As emotions are built into their definition of internalism, their understanding of internalism is certainly more qualified than UMI, which says all moral judgments, necessarily are motivating. But Corns and Cowan think that their objections apply to more qualified versions of internalism.³ One major goal of their paper is also to show that, even if Zagzebski or Prinz appeal to emotions, this approach still does not allow them to reach the conclusion they want to defend.

I do not attempt to defend UMI in a strict sense, either. Just like Zagzebski and Prinz, I want to defend a more qualified version of internalism, which I take to be reasonable. That is: in so far as the emotional dispositions that moral judgments consist of are manifested as actual emotions, the moral judgments will necessarily motivate the agents to act accordingly. This position will be fully developed in Section 4 where I give my positive account.

Another issue worth clarifying upfront is that internalism, whether UMI or a more qualified version, allows cases where competing motivation overrides the motivational force of a moral judgment. For example, Sally makes the judgment that it is morally wrong for her to cheat on her romantic partner. The judgment motivates her to stay away from any potential opportunities for her to be unfaithful. However, she has a strong desire to engage in a romantic relationship with her coworker Morgan which overrides her motivation to stay away from a potential affair. She ends up having an affair with Morgan. To put it another way, it is not a problem for internalism if one eventually does not act in accord with the moral judgments they make. It would be a problem for

³ For example, Section 2 of C&C's paper argues that their objections apply to conditional internalism as well.

internalism, however, if a moral judgment did not motivate one to act at all when other conditions are fulfilled.⁴

In the rest of this section, I will introduce Zagzebski's and Prinz's arguments. Their general strategies for defending internalism are similar, as they both argue that emotions play some constitutive role in certain types of moral judgments and that emotions have motivational power.

2.1 Zagzebski's Pluralist Approach

Zagzebski argues that a central group of moral judgments, which she calls "ground-level moral judgment" are expressions of emotions, and they are propositional in form (2003, 108). For example, we regularly make moral judgments such as—"What he just did was rude," "It was very kind of you to do so and so," or "This act is contemptible." These are the moral judgments that we make here and now (or in Prinz's terms, "occurrently") when confronting the intentional objects of our judgments. When we make these judgments, we are expressing our emotions regarding these intentional objects. Zagzebski demonstrates this point by introducing the idea of "thick affective concepts". In the examples above, the judgments use the concepts of "rude," "kind," and "contemptible," which are all thick affective concepts. They are thick in the sense that they have specific descriptive content. In contrast, concepts like good, bad, right, or wrong are considered thin concepts because they do not have specific descriptive contents and can be more generally applied. For example, rude can only apply to certain types of behavior and we can also say that rude behaviors are all bad, but not everything bad can be called rude. In addition to having descriptive contents, thick affective concepts are affective as the agents must be in distinctive

⁴ I recognize that it could be difficult to figure out whether one is motivated by a moral judgment in real life. But to the very least, we could make inferences from a person's behavior. For example, Sally might eventually cheat if competing desires end up gaining control. But if she is indeed motivated by her moral judgment not to cheat on her partner, we can expect her to act in certain ways. For example, she might try to resist the competing desires in the first place by avoiding Morgan or trying to talk herself out of the thought of sleeping with Morgan. She will also experience guilt after having the affair with Morgan.

emotional states when applying such concepts (114-115). We must have the feeling of being offended when we see something as “rude”. We must feel pity to see something as “pitiful”. We must feel contempt to see something as “contemptible”. Hence, in making these ground-level moral judgments that employ thick affective concepts, we must be in emotional states. We also express the emotions we experience through moral judgments using these thick affective concepts.

But as Zagzebski points out, not all moral judgments are made here and now like these ground-level judgments where we fully invoke our emotions. Sometimes we use thick affective concepts to make moral judgments about hypothetical situations, incidences that happened to others, or events that took place in the past, which she calls Level 2 judgments (2003, 119-120). A person might say “X did a contemptible thing to Y” to refer to something she hears about but did not witness. In these scenarios, even though we are still using the thick affective concepts, we might not invoke full-fledged emotions. At most, we invoke “a faint copy” of the memory about how these emotions feel (2003, 121).

As we move to what Zagzebski calls Level 3 moral judgments, we start using concepts like “*right, wrong, duty, should, ought, and good*” that are more abstract (2003, 121). In these cases, the focus shifts from the intentional objects of our emotional responses to our actual responses. For example, in judgments like “It is morally *right* for me to help a person in need” or “I have a moral *duty* to respect other people’s choices regarding their private lives”, the concepts of “right” and “duty” are applied to the specific choices of actions made by the agents. We need not invoke emotions at all to be able to make these moral judgments. Imagine in a philosophy class, a student who believes in the harm principle says that she thinks we should not interfere with others’ life unless they harm someone else. She need not invoke emotions at all to make that moral judgment.

To be clear, though not required, one could invoke emotions when making such moral judgments. For example, Zagzebski points out that those people who have developed an emotional response to concepts like duty or right or wrong can make Level 3 judgments emotionally (2003, 122). Moreover, our understanding of Level 3 judgments is derivative from our understanding of Level 1 and Level 2 judgments. Part of the function of these thin concepts is to help us link thick affective concepts like “pity” with our practical actions in response to the sufferings of others. For example, when we make the ground-level judgment that “He is pitiful” when seeing someone was mugged, we might subsequently make a Level 3 judgment that “I morally ought to help him” which connects our emotions of pity to actions. As Level 3 judgments are grounded in ground-level judgments that are intrinsically emotional (as they express emotions), it will not be surprising that we often make Level 3 judgments emotionally.

Thus far, Zagzebski has argued for the intrinsic connection between emotions and the making of moral judgments. The ground-level judgments are necessarily emotional and the abstract Level 3 judgments, if directly derived from these ground-level judgments, are also going to be emotional. Importantly, for Zagzebski, these moral judgments are not just emotional but that they are expressions of emotions, and in this sense, emotions are constitutive of these judgments.

Next, in the second part of her argument for internalism, Zagzebski maintains that emotions are motivating. Here she does not give a full-blown analysis but still provides important insights. She thinks emotions are motivating because “they combine affectivity and intentionality”, which differentiates them from affective states that do not seem to have intentional objects, such as sensations (2003, 115).⁵ Importantly, intentional objects give emotions “something specific in the

⁵ Zagzebski has not explicitly stated what she means by affective or affectivity, which I think has led to the false impression that she might be talking about affectivity in the sense that Corns and Cowan understands, *i.e.*, (un)pleasantness. I will return to this point in 3.1 where I argue that she is not relying on pleasantness/unpleasantness of emotions to argue that emotions motivate action.

world around us towards which affect is directed and which leads us to respond in ways characteristic of the emotion” (2003, 115). She also references Frijda’s account of emotions to emphasize that emotions are action dispositions in the sense that they involve a tendency to “change the world” in ways characteristic of the given emotion. For example, compassion, as noted by Zagzebski, involves a tendency to alleviate suffering (2003, 116). In Section 4, I return to Frijda’s account with more detail to explain the point that emotions involve action tendencies.

The last point worth emphasizing is that Zagzebski’s account allows us to explain why the motivational force of moral judgments might be strong in certain circumstances and weaker in others. Ground-level moral judgments such as “You are being rude” are made in emotional states and express the emotion of anger, so they are always motivating. But as we move away from these central cases to more abstract moral judgments, emotions might become weaker and so does the motivational power of the moral judgments. Sometimes judgments might not be emotional and hence not motivating at all. I read in the local newspaper that someone was brutally beaten up just for revenge over a minor squabble. Outraged by what the perpetrator did and empathetic towards the victim, I conclude that one should not use violence against others in such circumstances. I will be motivated to refrain from using violence in similar situations. On a different occasion, I might make a more abstract judgment that one should not use violence for unjustified reasons, invoking no emotions. The latter moral judgment will not motivate me to do any specific things.

In short, Zagzebski’s account of moral judgments and emotions supports a version of internalism that applies to moral judgments that are expressions of our emotions in response to situations in life. If a moral judgment expresses an emotion, then the judgment itself will be motivating. The stronger the emotion is, the stronger the motivational force will be.

2.2 Prinz's Sentimentalist Approach

Prinz makes an argument for internalism that is very similar to Zagzebski's:

“P1. Moral judgments consist of emotional attitudes.

P2. Emotional attitudes are motivating.

C. Therefore, moral judgments are motivating” (2015, 70).

In this argument, Prinz focuses on a specific type of moral judgment—moral judgments that use thin concepts such as good or bad, moral or immoral, and right or wrong and moral judgments that concern obligations from a first-person perspective (2015, 64 and 69). Some examples include “killing an innocent person is *wrong*” and “I *should* donate to Oxfam.” It is no surprise that Prinz focuses on this particular group of moral judgments because what puzzles us in the motivational internalism debate is why people sometimes appear to be *wholly* unmotivated *at all* to act in accordance with the prescriptions of these first-person action-guiding moral judgments. Notice that these first-person action-guiding moral judgments coincide with Zagzebski's Level 3 moral judgments, a point that I will return to shortly. But first, I will quickly explain what Prinz means by P1.

In P1, Prinz argues that first-person action-guiding moral judgments consist of emotional attitudes. Here, an emotional attitude means an emotion directed toward a target. For example, the judgment that “stealing is wrong” consists of a negative emotion toward stealing (2015, 70). He thinks that there is a constitutive relationship between emotions and moral judgments because moral judgments express our underlying emotional dispositions (2006). This view can be better understood through his sentimentalist approach to morality. As a sentimentalist, Prinz thinks that “to believe something is morally wrong (right) is to have a sentiment of disapprobation (approbation) towards it”, and the sentiment just is “a disposition to have emotions” (2006, 33-

34). In other words, emotional dispositions are essential to having moral judgments. To understand what Prinz means by emotional dispositions (or dispositions to have emotions), it is helpful to first compare his view on moral judgments with Zagzebski's.

As said, first-person action-guiding moral judgments are just Level 3 judgments under Zagzebski's taxonomy, and she thinks these moral judgments tend to be less emotional or even not emotional at all each time the judgments are made. It might look like Prinz is saying that Level 3 judgments must be emotional, which is different from Zagzebski's view. But they are not having a substantive disagreement. Zagzebski thinks of emotions in terms of emotions that are actually manifested. Prinz considers *dispositions* to have emotions. Such dispositions might remain dormant and only develop into actual emotional states in response to certain external conditions. For example, we might have the disposition to have anger towards the torturing of innocent people. Once we encounter a concrete case of torturing, the disposition will be manifested into actual occurrences of emotions such as anger or disgust towards the torturing. So, for Prinz, to have a moral judgment that torturing is wrong is to have the disposition to have emotions of some sort towards torturing, such as indignation, disgust, outrage, resentment, *etc.* When he says that first-person action-guiding moral judgments consist of emotional attitudes, it could be understood as saying that having these moral judgments require us to have dispositions to have certain emotions towards some objects.

In other words, for both Zagzebski and Prinz, one can make a moral judgment without being in an ongoing emotional episode. Zagzebski would accept this statement because she thinks that first-person action-guiding judgments (Level 3 judgments) do not require the occurrence of emotions. Prinz would accept the statement so long as the person has a disposition that will be manifested into actual occurrences of emotional episodes when circumstances rise. As he

succinctly puts it, “[e]xperienced emotions serve as a sincerity condition” (2006, 38). He further gives an interesting example—“if I am never outraged by gender discrimination, I am paying lip service to equity” (2006, 38).

So far, we have seen that Prinz thinks that emotions play constitutive roles in moral judgments because moral judgments express our underlying dispositions to have certain emotions which will be manifested under the right eliciting situation.

Next, the second premise of Prinz’s argument says having emotions toward something motivates the agent to act. Prinz cashes out the concept of motivation as states that “vie for practical control” (2015, 63-65). In other words, a mental state can be counted as motivational if it competes against other mental states in exerting influence over our actions. Along this line, psychological states such as sex drive, hunger, thirst, and desires can all be categorized as motivational states (2015, 63)t

Prinz thinks that emotions certainly compete with other motivational states to influence our actions. For example, emotions such as anger, contempt, disgust, rage, annoyance, *etc.* can affect our actions as they “promote avoidance, ceasing, intervention, withdrawal, and when anticipated, preventative measures” (2006, 34 and 36). Although Prinz does not explicitly invoke his own theory of emotions in the article where he defends internalism, looking at his own theory of emotions might help us better understand why he thinks that emotions “vie for practical control”. In following the Jamesian tradition, Prinz’s theory of emotions foregrounds the importance of bodily changes. For example, fear involves a racing heart and other physiological changes (2004, 69). But a shift in Prinz’s theory of emotions is worthy of highlighting.

In his earlier work, Prinz thinks that emotions are embodied appraisals, which just means that emotions are perceptions of our bodily changes. Here Prinz also adopts a teleosemantic

theory of representation to explain how emotions represent things in the world. According to the teleosemantic account, a state represents X by its function of being reliably caused by X. For example, fear is the perception of certain patterned bodily changes, which reliably occur when we are in danger. We can say fear represents danger, as it has the function of being reliably caused by danger.⁶ Additionally, emotions also have what he calls “valence markers” (2004, 163). Our brains assign emotions like happiness, contentment, and calmness with positive valence markers (2004, 162-173). Similarly, our brains assign emotions like annoyance, fear, sadness, and distress with negative valence markers (2004, 162-173). But more importantly, these valence markers serve as inner reinforcers that tell us to fulfill certain inner goals which eventually lead us to perform certain actions. For instance, fear has a negative valence that tells us we want “Less of this [bodily feeling]!” which leads us to perform actions such as fleeing away which can help us get rid of the feeling of fear. Similarly, positive valence markers of emotions like happiness tell us we want more of the same feelings which leads us to further engage with the world to get more of the same feeling (2004, 174). The joy of having one’s favorite food tells us “More of this!”, leading us to eat more of the food (Prinz, 174). In short, in his earlier work, he thinks that our physiological bodily changes allow emotions to represent things in the world and that the connection between emotion and action is indirectly modulated through the positive or negative valence markers that our brains assign to our internal bodily states.

Since then, there has been a shift in Prinz’s view on emotions. Specifically, his focus shifts from valence to the action aspect of emotions. In a co-authored article, Prinz and Daniel Shargel propose an enactivist account of emotions which focuses on the potential for actions brought forth

⁶ In his earlier book *Gut Reactions*, Prinz argues that emotions are not motivations, but motives that provide reasons for actions (2004, 193-194; 2015, 63).

by bodily changes when we are in emotional states (Shargel and Prinz 2018). According to this account, the connection between emotion and action is no longer due to the inner reinforcer of valence markers, but because changes in our bodily states prepare us to act. Shargel and Prinz argue that we are “*eo ipso* pulled [by the bodily changes] toward certain lines of behavior” (2018, 119). Motivation in this case is “immanent,” in the sense that “it is not merely an injunction to act” but rather, our bodies “push and pull us” to act (2018, 119).

This shift of focus from the valence to the potential of actions brought by bodily changes has already been foreshadowed in the 2015 article where he defends internalism. Just like Zagzebski, Prinz also defends the claim that emotions motivate actions by drawing from Frijda’s account of emotions to suggest that emotions are characteristically associated with action tendencies. Frijda’s account of emotions as action tendencies is referenced again in the 2018 article where he and Shargel propose enactivism. The idea of action tendency is that emotions activate our bodies to get ready to execute certain actions that help us maintain or achieve a certain kind of relationship with the external environment (Frijda 1986, 75). Again, Frijda’s account will be taken up with more details in Section 4. But the key takeaway here is that when defending internalism, Prinz is more on board with the idea of seeing emotions as action tendencies rather than his earlier view that relies on the inner reinforcer, valence marker, to explain how emotions motivate actions.

As will be discussed in more detail in the next two sections, C&C’s critique of Prinz’s defense of internalism focuses exclusively on his earlier views, which possibly leads to a misread of the argument for internalism that Prinz puts forward in the 2015 article.

To summarize the position of emotion-based accounts, Zagzebski and Prinz adopt similar strategies by first arguing that some moral judgments are constituted by emotions (Zagzebski) or that all moral judgments are constituted by emotional dispositions that will be manifested as actual

emotions when circumstances arise (Prinz). As the second step, they both argue that emotions are motivating (P2). Hence, for Zagzebski, a qualified version of internalism is true—if a moral judgment is constituted by emotions, then it will be motivating. For Prinz, moral judgments, in so far as they are genuine rather than mere lip service, will eventually motivate actions when emotional dispositions are manifested.

In the next section, I first will walk through three major objections that C&C raise against these emotion-based accounts for internalism.

3 CORN & COWAN'S OBJECTIONS

C&C raise objections to “the affective appeal,” by which they refer to “the attempt to vindicate a more empirically respectable and less theoretically controversial Internalism by appealing to affective episodes” (2021, 74). They further clarify that their targets are arguments that “appeal to the *affective aspect* of affective mental episodes” rather than those that appeal to other aspects of affective mental episodes (2021, 74). Here, affective mental episodes refer to conscious states that are “paradigmatically pleasant or unpleasant” which include things like emotions and sensations such as pleasure and pain (C&C 2021, 73; Corns 2014, 240).

Notice that unpleasant or pleasant mental states are often considered “valenced” states in the literature of emotion (*e.g.*, Teroni 2019). However, it seems that C&C refrain from calling the pleasant and unpleasant mental states “valenced” states on purpose, possibly because they think positive and negative valence markers might not map onto the pleasantness or unpleasantness of emotions.⁷

⁷ Corns recognizes that “valence” can overlap with pleasant or unpleasant feelings, as she writes in a 2014 article, “if Prinz is correct about valenced emotions, then there is a component that all negative emotions and painfulness share: aversive valence.” However, Corns and Cowan do not think valence is the same as affect, which they exclusively refer to feelings of pleasantness or unpleasantness, which are usually thought to be realized in human

So, how exactly are Prinz and Zagzebski's accounts problematic, according to C&C?

Recall

Unconditional Motivational Internalism (UMI): Necessarily, if *S* judges that she morally ought to ϕ , then *S* is motivated (at least somewhat) to act in accordance with her judgment. (2021, 72).

C&C argue that Zagzebski and Prinz's accounts cannot help them defend UMI. Specifically, they focus on the second premise which says that emotions motivate agents to act. They think that in order to defend UMI, which is a claim about necessity, the second premise of the argument needs to be a claim about necessity, too. That is, in their second premise, Prinz and Zagzebski need to argue that emotions are *necessarily* motivating in order to defend UMI. However, C&C think it is false that emotions necessarily motivate, because mental states that are pleasant or unpleasant can doubly dissociate from motivational states. What they mean by double dissociation is that one can be in an affective mental state without being in a motivational state, and vice versa. C&C argues that the dissociation takes place at both the personal level and the sub-personal level. By the personal level, they refer to states or processes of a person as a whole, and by the sub-personal level, they refer to states or processes that are parts of a person. For example, if I choose to buy a cup of coffee, it can be explained by my desire to drink coffee, which is a personal-level explanation. My action of purchasing coffee can also be explained by the firing of my motor neurons, which is a sub-personal level explanation.

C&C first argue that, at the sub-personal level, the neurophysiological systems that process our positive hedonics (dubbed 'liking') do not necessarily co-activate with the system that processes our motivation (dubbed 'wanting') (C&C 2021, 78; Robinson and Berridge 1993;

beings through the dopaminergic system. In the same 2014 paper, Corns explicitly states that "[i]t is again helpful to make clear what this feature—aversive valence—is not. I will later argue that it dissociates from unpleasantness."

Robinson *et al.* 2015; Pool *et al.* 2016). Here they are assuming that the pleasantness and unpleasantness of emotions roughly maps onto the ‘liking’ system and our motivations can be roughly mapped onto the ‘wanting’ system. Then they point out that these two systems are distinct—they involve different “neurochemicals, pathways, and structures” and they can also be “independently modulated” (C&C 2021, 78). Call this objection ‘Liking’ vs ‘Wanting’ objection, which I will specifically address in Section 4.1.

To support their objection, C&C cite studies from neuroscientific studies. For example, it was found that certain highly addictive substances such as nicotine do not produce much euphoria, but they are exceedingly ‘wanted’ (Robinson *et al.* 2015, 111), in which case one ‘wants’ something without ‘liking’ it. One might also ‘like’ something without ‘wanting’ it. In the treatment of drug addiction, the medical substance that completely blocks dopamine function reduces the craving for drugs, but when patients are asked to rate the level of pleasure brought by addictive drugs such as amphetamine or methamphetamine after the treatment, their subjective ratings do not change (Robinson *et al.* 2015, 111).

C&C tie the distinction between ‘wanting’ and ‘liking’ with the internalism debate, arguing that the dissociation between the neurophysiological systems of ‘wanting’ and ‘liking’ suggests that there is also a dissociation between affective mental states and motivations. In other words, emotions, by virtue of being pleasant or unpleasant, can happen without the occurrence of motivations.

C&C further propose that there could be dissociation at the personal level between affective mental states and motivation as well. They offer two examples to show that we can have emotions towards something without feeling motivated concerning it.

“Unwarranted Guilt: If someone knows that they’re prone to feeling unwarranted guilt after nights out, then they may be unmotivated by guilty feelings experienced on subsequent occasions” (C&C 2021, 81).

“Moral Despair: Having been over-exposed to horrible news, someone morally despairs about the world and thereby lacks motivation to engage in activism” (C&C 2021, 81-82).

In both cases, although the agents are in emotional episodes, they do not seem to motivate the agents to act. In the case of Unwarranted Guilt, the person continues to go out at night, despite having a sense of guilt, because they know that the feeling of guilt is totally unwarranted since it is only “chemically-induced” and “going out is not actually wrong” (C&C 2021, 81). This example shows that having emotions like guilt sometimes will not motivate us to act because we can completely ignore it by recognizing it is unwarranted. I call this counterexample the Unwarranted Guilt objection.

In the case of moral despair, the person refrains from actions associated with activism because they are in the grip of despair. This example is meant to show that despair characteristically does the opposite of what internalists argue: they predispose us to refrain from acting rather than acting. Call this counterexample the Moral Despair objection. It will be addressed in Section 5 after I give my positive account.

In Section 4, I will first address the ‘Liking’ vs ‘Wanting’ objection. I want to first make the case that the Liking vs Wanting dissociation, although likely to be true at the neurophysiological level, does not mean we have to reject emotion-based internalism. The reason is that existing emotion-based accounts are not primarily appealing to affect to defend internalism. More importantly, the right way to cash out the agents’ motivations when they are in emotional states is through action tendencies. I will incorporate this idea with the existing emotion-based

arguments of Zagzebski and Prinz to propose a revised argument. My argument will be able to explain how emotions motivate actions and, more importantly, why we would be motivated by moral judgments to act in accordance with what they prescribe, an important element missing from Zagzebski and Prinz's account. Filling the gap also allows us to better understand the two examples raised in the Unwarranted Guilt objection and the Moral Despair objection, which I will respond to in Section 5.

4 EMOTIONS ARE FUNCTIONALLY MOTIVATING

In the 'Liking' vs 'Wanting' objection, C&C pick out a particular aspect of emotions, the pleasantness or unpleasantness of emotions, or as they put it, the affective aspect of emotions. They argue that emotions, in so far as they are affective mental states, do not necessarily co-activate with our motivational systems. However, when defending internalism, neither Zagzebski nor Prinz says that emotions are motivating only or primarily by virtue of being pleasant or unpleasant. In this section, I argue for two conclusions. First, emotion-based accounts do not primarily appeal to the affect of emotions to make the case for internalism. Therefore, the 'Liking' vs 'Wanting' objection simply does not apply to emotion-based accounts. Second, emotion-based internalists also do not need to argue that the pleasantness or unpleasantness of mental states is the source of moral motivation in order to defend internalism, because emotions are functionally motivating—they serve the function of preparing us for actions through changes in our bodily states and these actions help fulfill certain goals in reaction to certain external conditions.

4.1 Response to 'Liking' vs 'Wanting' Objection

To defend the first claim, I look at how Zagzebski's and Prinz's accounts deal with affect. Again, by affect, C&C refer to whether the emotion feels pleasant or unpleasant.

The way in which Zagzebski uses the word “affect” might have misled C&C to think that she is narrowly appealing to the pleasantness or unpleasantness of emotional episodes. As mentioned in Section 3, C&C quote Zagzebski when she says that the “affective aspect [of emotions] is ‘pushy’ ” (C&C 2021, 77; Zagzebski 2003, 116). However, Zagzebski seems to use the term “affect” and the term “emotion” interchangeably throughout her paper. When talking about “thick affective concepts” or affective aspects of emotions, Zagzebski might use the word “affective” to represent features of mental states that make them emotional, such as phenomenal feelings characteristic of each emotion, emotion’s control over the way we think or act, *etc.* She is certainly not thinking about affect in terms of likes/dislikes or pleasantness/unpleasantness.

To put it another way, the debate between C&C and emotion-based internalists is possibly a verbal dispute. If we replace the word ‘affect’ with pleasant or unpleasant feelings, then Zagzebski will agree with C&C that it is not the case that emotions are motivating only because they have pleasant or unpleasant feelings. As touched on in Section 2.1, for Zagzebski, intentional objects are also an important part of the story about why emotions are motivating. She thinks that emotions are motivating because “they combine affectivity and intentionality”, which differentiates them from affective states that do not seem to have intentional objects, such as sensations (2003, 115). In her own words, intentional objects “giv[e] [emotions] something to which the agent is motivated to respond” (2003, 116). Moreover, she is clearly aware that emotions should be characterized through the lens of action tendencies. For example, as mentioned in Section 2.1, Zagzebski emphasizes that emotions should be understood as action dispositions that prompt us “to change the world in ways characteristic of the given emotion” (2003, 116). Hence, a holistic reading of Zagzebski’s account suggests that she is not saying that emotions are

motivating only or primarily because of the pleasantness or unpleasantness of emotions, but that emotions are motivating in the sense that they dispose us to act in certain ways.

Similarly, Prinz's argument does not attribute the source of motivation only to the pleasantness or unpleasantness of emotions. Admittedly, in his earlier work, valence plays the role that links emotions to actions, and the concept of valence is taken to be closely related to the pleasantness or unpleasantness of emotions in the literature of emotion.⁸ But even so, Prinz rejects the view that emotions with positive valence are pleasant and that emotions with negative valence are unpleasant (2003, 167-168). Rather, as explained earlier, he thinks of valence as inner reinforcers that tell us whether we want more of a bodily feeling (2003, 173-174).

Moreover, it might be a more accurate reading of his view if we understand his argument for internalism (published in 2015) in light of the article co-authored with Shargel published only three years later. In this 2018 paper, he rejects the earlier approach of using valence to explain how emotions lead to actions and instead endorses an enactivist account of emotions. According to the enactivist account, the potential for actions is brought forth by bodily changes when we are in emotional states. It is through bodily changes that we create more potential for action. For example, as they mentioned, when a swimmer realizes that she is too far away from the shore, she experiences fear. In fear, her body "mobilizes more forcefully to support strenuous action, helping her return safely" (Shargel and Prinz 2018, 120). Prior to fear, the sea was not a place to be escaped, and it usually does not seem to be a possible task for the swimmer to swim across such as distance. Only through bodily changes that fear brings does the possibility of strenuously escaping the sea become available (2018, 120).

⁸ *E.g.*, Fabrice Teroni writes in a 2019 article that "Let us call all valenced states 'affective'. This is not meant to be stipulative – we pre-theoretically think that these states form a family, and this is surely due to the fact that they are valenced." (Teroni 2019, 103)

Additionally, even in the 2015 article in which Prinz defends internalism, there were signs suggesting that his view has already shifted to focus more on the action tendencies of emotions. For example, in defending the second premise of his argument which says that emotions motivate, he references the work of Frijda and says that “each basic emotion is associated with a characteristic action tendency” (Prinz 2015, 74; Frijda 1986). He goes on to give a list of basic emotions to illustrate this point: “anger promotes [the action tendency of] aggression, fear promotes [the action tendency of] flight, and disgust promotes [the action tendency of] withdrawal of the senses” (Prinz 2015, 74).

To sum up, although it might seem to be the case that emotion-based accounts proposed by Zagzebski and Prinz appeal to the affect of emotions to make the case that emotions are motivating, there are good reasons to think that this interpretation of their views is misleading. Neither of them relies on the pleasantness or unpleasantness of emotions to argue that emotions motivate actions. Therefore, even if ‘liking’ can doubly dissociate with ‘wanting’, the ‘Liking’ vs ‘Wanting’ objection does not apply to their accounts.

4.2 The Missing Piece of the Puzzle

More importantly, emotion-based internalists do not have to rely on hedonics of emotions to explain why emotions motivate actions, which is the second claim that will be defended in this section. Rather than hedonics, emotions motivate because of their functional role of causing actions that fulfill the goal of the emotions. Before moving to the positive account, I want to point out a gap in the existing emotion-based accounts for internalism.

Prinz and Zagzebski have already pointed out that emotions motivate us to act because emotions involve tendencies to act in ways characteristic of each emotion. But they both miss an important piece of the puzzle, which is how an emotion towards something that happens in the

world motivates the agent to act in a way that is in accordance with the prescription of the agent's moral judgments.

I propose a revised argument to improve on the existing emotion-based arguments, which will fill in the gap. The improved argument will also be able to address the Moral Despair and the Unwarranted Guilt objections that C&C makes against Zagzebski and Prinz.

4.3 The Revised Argument

To illustrate why seeing emotions as functional states makes a stronger case for emotion-based internalism, I propose the following revised argument that builds on Prinz's and Zagzebski's arguments.

P1. If S makes a genuine moral judgment that she morally ought to ϕ , the moral judgment necessarily consists of a disposition to have an emotion about a specific intentional object.

(Sentimentalism)

P2. A disposition to have an emotion about a specific intentional object, if manifested as actual emotion, will necessarily motivate the agent to take actions that help achieve the goal of the emotion.

(MT)

P3. Therefore, if S makes a genuine moral judgment that she morally ought to ϕ and if the emotional disposition that constitutes the moral judgment is manifested as actual emotion, the emotion will necessarily motivate S to take actions that help achieve the goal of the emotion (from 1 and 2).

P4. If S makes a genuine moral judgment that she morally ought to ϕ , and if the emotional disposition that constitutes the moral judgment is manifested as actual emotion, ϕ ing will necessarily help achieve the goal of the emotion.

C. Therefore, necessarily, if S makes a genuine moral judgment that she morally ought to ϕ and if the emotional disposition that constitutes the moral judgment is manifested as actual emotion, S is motivated to ϕ .

This argument improves on the existing emotion-based arguments by spelling out how an emotion that responds to something in the world motivates the agent to act in a way that is in accordance with what the agents' moral judgments prescribe them to do. I am aware that a strict reading of UMI is implausible, but a qualified version is reasonable: in so far as the constituting dispositions to have emotions are actually manifested as occurrent emotions, the moral judgment will necessarily motivate the agents to act accordingly.

In the next three sections, I will go through each of the premises to defend the above argument.

4.4 P1: Aboutness and Specific Objects

P1 says if S makes a genuine moral judgment that she morally ought to ϕ , the moral judgment necessarily consists of a disposition to have an emotion and such emotion must have specific intentional objects.⁹ In saying that moral judgments consist of dispositions to have certain kinds of emotions, I am on board with the view endorsed by Prinz that moral judgments are expressions of underlying emotional dispositions. But different from the original first premise, the revised P1 specifies that such emotions must have specific intentional objects. In other words, emotions that can constitute moral judgments are the ones that respond to specific intentional objects.

⁹ If one subscribes to Zagzebski's pluralist view, they can modify P1 to say that when emotions are invoked, moral judgments that one morally ought to ϕ necessarily consists of an emotion directed towards an intentional object O. Their conclusion will accordingly become the following: necessarily, if S makes the moral judgment that she morally ought to ϕ and the judgement consists of emotions, then S is motivated to ϕ . This conclusion is still in line with an emotion-based defense for internalism.

In the discussion about “thick affective concepts”, Zagzebski already gives us a simple account of how emotions can be about an intentional object: some evaluative process must also be involved when we apply emotionally charged words. She is certainly on the right track in thinking that emotions are evaluative. I appeal to Richard Lazarus’s influential appraisal theory of emotions to explain how emotions have intentionality. The appraisal theory holds that emotions involve an appraisal of core relational themes (Lazarus 1991). Core relational themes refer to things in the environment that cause certain reactions in agents. Fear involves an appraisal of danger in our surroundings. Anger involves an appraisal of offense that we encounter. Sadness involves an appraisal of loss in our lives. Guilt involves an appraisal of our own moral transgression (Yip 2022, 861; Scarantino 2014, 181). Moreover, I endorse a teleosemantic theory of representation: an emotion E is about X if it has the function of being reliably elicited by X. Here, the reliable elicitation of emotion by a core relational theme is another way of saying there is a robust covariance relationship between the emotion and the core relational theme. Fear is about danger because it has the function of being reliably caused by danger. Anger is about offense because it is reliably caused when someone offends us. Guilt is reliably elicited when we recognize our own moral mistakes. Hence, emotions are about these core relational themes.¹⁰

So far, I have maintained that, on the basis of existing emotion-based accounts, I endorse the view that moral judgments necessarily consist of dispositions to have emotions in the sense that they are expressions of such emotional dispositions. I also just argued that emotions are about

¹⁰ I recognize that in Prinz and Shargel 2018, Prinz modifies his view on how emotions represent things by adopting an enactivist account of emotions. For example, the enactive content of fear is no longer the danger. Rather, it is the situation that presents itself as something that one should escape from. I think the core relational theme is the right way to understand how emotions relate to things in the environment. But even if one has a different view on whether emotions have content or how it represents things in the world, one can still rely on the form of the argument that I propose in this section. One can modify the part on “intentional objects” in premise 1 and premise 2 to adjust to their own theories of emotions. *See*, Shargel and Prinz 2018. *See also*, Scarantino and de Sousa, 2021.

core relational themes. If these two claims are true, we can conclude that moral judgments necessarily consist of dispositions to have emotions about core relational themes.

Yet, that is not enough. P1 argues that first-person action-guiding moral judgments consist of dispositions to have emotions about specific rather than more general objects. The key here is that the underlying emotions necessarily are about specific intentional objects.

To start with, our moral judgments are often about specific scenarios even in cases where they are not first personal. We might judge that it was wrong for our friend's ex-partner to cheat on them. The emotions expressed by this moral judgment, be it disgust, anger, or a mix of both, are about the specific action of another person: the cheating of the ex-partner of our friend. A stranger offered to help as we struggled to put our luggage into the overhead bin when boarding a plane. We judged that it was morally supererogatory for the stranger to give an extra hand. In so far as this judgment was constituted by emotions, it would consist of emotions that approve what the stranger did, which could be gratitude or other similar emotions. Regardless of what exactly the emotion was like, it must be *of* about the help that the stranger offered.

Similarly, first-person action-guiding moral judgments are about specific things in the world. Again, first-person action-guiding judgments use concepts such as ought, right/wrong, good, *etc.* to make moral judgments from the perspective of "me (*e.g.*, "It is morally right for me to save a drowning kid" or "I morally ought to stay faithful to my partner.") Seeing how much our friend was traumatized by the cheating ex-partner, we judged that it would be wrong for us to cheat on our own partners, too. In practice, this judgment can be constituted by various emotions such as disgust, anger, a blend of both, or some other similar emotions, depending on the actual scenario. Whichever the emotion is, it is necessarily about the action of cheating, or more precisely, a hypothetical situation where we cheat on our partners.

Of course, some emotions do not seem to be about anything specific. For example, one might get up feeling quite joyful or one might get up feeling grumpy or sad. But none of these emotions can be expressed through moral judgments because these emotions are not about anything of moral significance. In other words, I am not claiming that emotions are always about specific objects. But in the case where emotions do constitute a moral judgment, they are necessarily about specific objects.

One might still disagree by arguing that the emotions that give rise to our moral judgments could be about things in general. For example, one who lives in an autocracy might judge that the surrounding political environment is morally horrifying and feels sad about the surrounding in general. But notice that even if one makes an overall evaluative claim about the environment that one lives in, it is still about a set of specific things. This person is feeling morally horrified exactly by all the institutions, incidences, and individuals that make up this environment, including propaganda by government-owned media, coercive policies, collusion between government officials and business oligarchs, repressive political parties in power, hypocritical politicians, people who support and help the regime to suppress others, *etc.* It is the summation of these specific things that the person is reacting to. Compare this scenario with an unlucky person who wakes up feeling sad or grumpy for no obvious reason. The former is still about specific things, although the “specific things” encompass a variety of things in the political environment. In contrast, the sadness, or the mild anger this unlucky person feels is not about specific things.

To conclude this discussion, although emotions can be about a general state of the world, emotions that moral judgments express necessarily are about specific things. This entails the truth of P1, which holds that for first-person action-guiding moral judgments, which are a subtype of

moral judgments, in so far as they express a disposition to have certain emotions, the emotions must be about specific intentional objects.

Now that the truth of the first premise has been defended, the next subsection moves on to defend the second premise.

4.5 P2: Emotions as Action Tendencies with Goals

P2 claims that when emotional dispositions are manifested as actual emotions about specific objects, these emotions will necessarily motivate agents to act in ways that help achieve the goals of the emotions. Here I rely on the work by emotion theorists, including but not exclusive to the Motivational Tradition (MT) of emotions to defend P2. MT theories argue that emotions are motivational states in the sense that they perform the function of causing behaviors aimed at achieving certain goals (Scarantino and de Sousa, 2021). I accept this conceptualization of motivation. In saying that emotions *motivate* the agent to act, I refer to the idea that emotions play the functional role of causing behaviors that aim at achieving certain goals relevant to the emotions.¹¹ But how exactly do emotions play such functional roles?

Frijda, an MT theorist, provides an answer to this question by pointing out that emotions are best characterized as *action tendencies with control precedence*. By “action tendencies,” he refers to the “state of readiness to execute a given kind of action” (Frijda 1986, 70-71). When we enter an emotional state, our bodies and mental states are disposed to act in ways that are characteristic of the emotion. For example, when in “fear,” as noted by Frijda, the agent has the action tendency of “avoidance” (Frijda 1986, 88).

¹¹ In adopting this functional approach to motivation, I reject the following view: for a mental state to count as motivational, the agent must feel motivated. Sometimes we could be wrong about whether something motivates us as conscious feelings are unreliable. When buying books, we might feel that we are motivated to acquire knowledge, while in fact we often are too lazy to read them after buying them.

Control precedence refers to the fact that when an agent is in an emotional state, the emotion has precedence in terms of control over the agent's cognitive functions, bodily states, and subsequent actions. As pointed out by Frijda, emotions tend to "interrupt other ongoing programs and actions," "clamor for attention," "preempt the information-processing facilities," and "persist in the face of interruptions" (Frijda 1986, 78). An example of a man who gets angry might be helpful to illustrate the control precedence that emotions have. After being scolded with offensive remarks, Andy becomes furious. Anger starts to gain global control over Andy's cognitive functions, bodily states, and subsequent actions. His attention is completely drawn to the situation. He stops doing whatever he was previously doing. His heartbeat rapidly increases. He starts to draw inferences that he usually doesn't make—"If a person makes an offensive remark, they should be punished with violence." He also starts to recruit past unpleasant memories related to the offending person which makes him more assured that this person deserves a punch. Meanwhile, his muscles become tense, and his hands claw. Having gone through all the above changes, Andy is now ready to confront the offending person.

One need not commit to a motivational account of emotions in order to recognize that emotions are typically associated with action tendencies. In fact, many theorists, including those outside of the motivational tradition, accept that emotions are associated with action tendencies that dispose us to act in ways typical of the emotion. For example, as mentioned in earlier sections, Prinz, who is not in the motivational camp, cites Frijda to make the case that emotions are associated with action tendencies (2015).

Of course, pointing out that emotions have action tendencies alone is not enough. To defend P2, which claims that emotions lead to actions that fulfill the goals of emotions, I still need to explain how emotions, which are about some objects in the world, can have goals. Again, MT

theories provide us with a plausible answer. Frijda thinks that because emotions involve action tendencies, they can be “defined by... end result[s] aimed at or achieved,” and the goals of emotions can be “inferred from behavior” of agents (1986, 70-71). A helpful way for understanding his claims is by looking at the flexibility of our behaviors when in an emotional state. For instance, when in fear, we might choose to hide, run away quickly from the danger, and/or ask for help along the way, all of which is to avoid the potential threat from the danger, such that we could protect ourselves. Hence, we could say that fear has the action tendency of avoidance to achieve the goal of self-protection (Frijda 1986, 88).

Furthermore, Andrea Scarantino, who is also a theorist from the MT tradition, gives an elaborate motivational theory of emotions, where he connects the intentionality of emotions and the goals of emotions together. He points out that emotions have two functions: (1) the function of representing core relation themes and (2) the function of achieving a certain relational goal (2014, 178). Think in terms of the example of fear again. Fear represents not only danger but also the goal of self-protection since fear involves the action tendency of avoidance which helps one hide from the source of danger to achieve safety.

Reactive attitudes, the emotional attitudes that we have in reaction to attitudes that others have towards us, as manifested in their actions, are also good examples to demonstrate Scarantino’s point (Strawson 1962). Emotions that count as reactive attitudes are all associated with action tendencies such that these emotions all functionally represent some facts about the world and cause behaviors to serve certain goals typically associated with the specific emotions. For example, moral outrage, which is directed towards some wrongful conduct, has the action tendency of attacking or removing the source of obstruction. This action tendency helps us achieve the goal of regaining control. Gratitude, which is typically directed towards those we think have

benefited us, is associated with the action tendency to express appreciation and do something nice in return, with the end goal of benefiting the recipient of this emotion reciprocally. Guilt, which is usually directed at our own moral transgression, is associated with the action tendency of doing things beneficial towards the intentional object of the emotion, *i.e.*, those we feel guilty towards. We might do various types of things to make up for our moral mistakes with the end goal of compensating those whom we feel guilty towards (Scarantino 2014, 181). Other important emotions, such as compassion that we have towards others are also motivating in the sense that they involve the action tendency of more engagement with the person in need including giving more care. Our voices tend to get tender, and our attention tends to get fixed on the person in need of help. Such an action tendency prioritizes actions that help fulfill the goal of helping others in the moral community.

I anticipate the objection that not every emotion motivates agents to act. For example, in the case where a person wakes up feeling happy or sad, the emotion does not seem to have a goal. Observe a person who is simply in joy or sadness in a global sense, *i.e.*, they are not happy or sad about specific things but are happy or sad in their reaction across things in the external world. It is unclear what goals they might have. My response is that when the person is engulfed in sadness in the global sense, sadness plays the functional role of causing behaviors of certain sorts, in this case, a general disengagement with the surrounding. Depending on the exact kind of sadness or despair the person feels, we might conclude differently about the goal of this type of emotion that promotes withdrawal behaviors. But one plausible interpretation of the goal is to restore oneself. In distancing oneself from the world, one rests, preserves energy, and restores oneself to a neutral position, ready to engage with the world again. Similarly, a person who is just happy about the world generally has the action tendency to engage more with the world. As put by Scarantino, this

joyful person is “ready to engage in an open range of actions”, and she “actively prepares for this open engagement with the world with a generalized state of arousal” (2015, 171). By interpreting the joyful person’s behavior, we can conclude that the goal of a general state of joy, plain as it sounds, is just to engage more with the world generally.

Even if one thinks that a person in joy or sadness in a global sense does not seem to be motivated to achieve any goals, P2 can still be left intact, because P2 and my overall argument focus on emotions with specific intentional objects. Specific intentional objects give emotions specific things to react to and accordingly more specific goals to fulfill.

So far, I argued that moral judgments necessarily consist of dispositions to have emotions about specific intentional objects (P1). As these emotions are about specific intentional objects, when they take place, they necessarily motivate the agents to act in ways that help achieve the goal of the emotions (P2). Hence follows P3, a logical necessity of P1 and P2, which essentially says that moral judgments consist of dispositions to have certain emotions and when these emotions are manifested, they necessarily cause the agents to act in ways that fulfill the goals of the emotions. However, up till this point, there still needs an explanation for the following question: why would the constitutive emotional disposition of the moral judgment (that one morally should ϕ), when manifested as actual emotions, necessarily motivate agents to act in accordance with the prescription of that judgment, *i.e.* to ϕ . The answer is that the action of ϕ ing is the specification of how exactly the goals of the emotions should be achieved, an idea I will elaborate in the next subsection.

4.6 P4: From Goals to Specific Actions

P4 states that if S judges that she morally ought to ϕ , the action of ϕ ing necessarily specifies what should be done to help achieve the goal of the emotion that the moral judgment consists of.

The key point here is that the action prescribed by the moral judgment is supposed to help advance the goal of the emotion, although one might be false in their judgment. For example, suppose that one forms the moral judgment that they morally ought to donate to charity and suppose the judgment consists of a disposition to have compassions towards those in need. Such emotions have the goal of helping others in the moral community. In forming the moral judgment, the agents specify for themselves the specific action that helps achieve the goal of the underlying emotion, which is donating to charity. Notice that the specified action and the goal of the emotion must be consistent with each other. Fear for charity, for example, cannot be the constitutive emotion of that same moral judgment, since the goal of the fear would be to avoid that charity, which is inconsistent with making donations. If one feels disgusted by charity and despise them for whatever reasons, it would be odd that they form the moral judgment that they morally ought to donate to charity.

But how exactly is the intentional action specified when we form first-person action-guiding judgments? Zagzebski makes a helpful note. She points out that what emotion disposes us to do is usually “vague and fairly unformed”, but “[a]ction requires something more specific to direct the agent to act in one way rather than another way that satisfies” what the emotion desires (2003, 116). She further points out that, “[s]ince the motivating aspect of emotion is relatively unfocused, it needs to be shaped by experience, knowledge, and an understanding of the entire context. In short, it needs to be shaped by practical reason” (2003, 116).

Scarantino makes a similar point by drawing from philosophy of action. He argues that emotions can have subgoals that help achieve the more abstract goal of the emotion. When we are in action, we use rational control to make sure that “the emotion’s relational goal is translated into

a set of sub-goals that is instrumentally adequate” in helping us achieve the relational goal (Scarantino 2014, 172).

To add to Zagzebski’s and Scarantino’s discussion here, I want to highlight that in the formation of first-person action-guiding moral judgments, we utilize contextual information and our practical reason to figure out what specific action we could do, while emotions give general guidance about what the action should ultimately aim at. Through this process, the goal of the emotion is instantiated by the specific action the moral judgment prescribes. Hence, a moral judgment that one morally ought to ϕ points us toward the sub-goal that necessarily helps reach the goal of the emotion.

By this point, I have defended the claim that our first-personal action-guiding moral judgments necessarily consist of dispositions of emotions and if such emotions are manifested, they will necessarily motivate the agent to act in ways that help achieve the goal of the emotions (P3). I have also shown that the action prescribed by these moral judgments helps achieve the goal of emotions that constitute them because the prescribed actions are specified subgoals (P4). It follows from these two premises that, necessarily, if S judges that she morally ought to ϕ , S is motivated to ϕ .

To better illustrate how this argument works, consider the following example. Recently, the firm that Jackie worked for decided to promote a junior staff to fill a vacant senior position. As an experienced senior staff, Jackie was put into the position of power to decide which junior faculty would be promoted. Lydia was obviously the most qualified employee as she was not only talented but also worked extremely hard. She already took work that was part of the job responsibilities of the vacant position, even though her title was just only a junior staff. Jackie, however, was personally fond of another candidate who was less qualified than Lydia. But as a person who values

fairness, Jackie found it despicable to take into consideration any personal feelings. Hence, she judged that she had a moral obligation to avoid evaluating candidates based on her personal ties with them, and she should instead evaluate them entirely based on their performance at work. To apply P1, Jackie's moral judgment against considering personal feelings consists of a disposition to feel disgusted (or other negative emotions) about a specific intentional object, *i.e.*, the counterfactual of her taking into consideration her personal ties to the candidates. As claimed by P2, a disposition to have an emotion about a specific intentional object, if manifested as actual emotion, will necessarily motivate the agent to take actions that help achieve the goal of the emotion. In the case of Jackie, her disgust about considering personal feelings would involve action tendencies to avoid the disgusting object, *i.e.*, to act unprofessionally and partially in this scenario. As maintained by P4, the specific action prescribed by moral judgment necessarily helps achieve the goal of the emotion. In this particular example, evaluating candidates entirely based on their performance at work would help serve the goal of avoiding the disgusting alternative—having her judgment tampered with personal affinity with one of the candidates. Hence, Jackie would be motivated to act in accordance with what was prescribed by her own moral judgment. We could even imagine her relying on a scoring sheet or other similar strategies to evaluate all the candidates based on their performance.

An advantage of the revised argument is that it allows us to explain why moral judgments, in virtue of being constituted in part by emotional disposition, can motivate us to do specific things they prescribe when the emotions are manifested. The next section will use the premises in the revised argument to explain away the two counterexamples that C&C raises against emotion-based arguments for internalism.

5 UNWARRANTED GUILT AND MORAL DESPAIR

In Section 4.1, I responded to the ‘Liking’ vs ‘Wanting’ objection. However, the other two objections—the Unwarranted Guilt objection and the Moral Despair objection—remain largely unaddressed. This section reviews these two objections in light of the argument proposed in Section 4.

5.1 Unwarranted Guilt

Recall the case of Unwarranted Guilt: “If someone knows that they’re prone to feeling unwarranted guilt after nights out, then they may be unmotivated by guilty feelings experienced on subsequent occasions” (C&C 2021, 81). The idea is that the person knows that “the affective episode is chemically-induced” and that going out at night “was not actually wrong,” so their subsequent actions are unaffected by the feelings at all. Crucially, C&C think that in this case the motivation is “cancelled” rather than defeated by other competing mental states (C&C 2021, 82). In other words, there is no motivational force whatsoever.

However, even looking at the case itself, it is false to say that the motivation of guilt is “cancelled”. The reason why guilt does not seem to affect the subsequent behavior of the person (they continue to go out at night) is precisely because it has been overridden by other considerations. The fact that this person, in anticipation of the guilt, has to plan in advance in order to ignore it only shows that the guilt does have a motivational force. Were they not to actively plan to disregard the guilt, this emotion would lead them to do something different, such as repenting. The emotion is therefore defeated rather than canceled. It is outweighed by other competing mental states, perhaps a determination to live one’s life without being interfered with by some guilt that is perhaps conditioned by an unwarranted moral authority. Or more simply, it is outweighed by a strong desire to go out and have fun.

Further, guilt motivates the agent by playing a causal in their subsequent actions. As argued in defense of P2, emotions should be understood as action tendencies that prepare the agents to act in certain ways to achieve the goal of the emotions. If one indeed feels guilty, it will involve bodily changes that actively prepare the agent to act in ways that fulfill the goal of guilt, which is to make up for the transgression. For example, one might have an action tendency to repent to the moral authority whom they feel guilt towards and have the thought that they should not repeat the same mistake ever again.

5.2 Moral Despair

Also recall the counterexample of Moral Despair: “Having been over-exposed to horrible news, someone morally despairs about the world and thereby lacks motivation to engage in activism” (C&C 2021, 82).

The key point that C&C want to show is that the person in moral despair (or depression, which is another word that C&C used to describe the case) is demotivated to do things, including engaging in activism.¹² There are a couple of ways to respond to this case, depending on how the detail of the case is specified.

One way to interpret the situation is that the person “morally despairs about the world” in a general way. If that is the case, then the morally despaired person is very similar to a person who wakes up feeling sad— they both have a low level of arousal, and the emotions are about the external world in general rather than anything in particular.

As argued in Section 3.4 where I defend P1, for an emotion to be possible to be expressed through first-person action-guiding moral judgments, it has to be about specific things. In this case,

¹² In their discussion of this example, C&C use the words depression and despair interchangeably. I assume that by depression, they do not mean chronic illness of depression, because it is not an emotion. I understand despair or depression in this case as something similar to sadness, with an additional element of hopelessness.

moral despair is not about specific things but about the surrounding in general. They cannot be the building block of a first-person action-guiding judgment to start with. In other words, whether this emotion is motivating or not does not affect the revised argument. So, this case is simply irrelevant to the debate about emotion-based internalism.

Still, I want to clarify that I think moral despair does motivate actions. As argued in defense of P2, the right way to think about emotions is through action tendencies. In this case, a person in moral despair has the action tendency to reduce engagement with the external world. Or in other words, to withdraw from the world. One plausible interpretation of the functional goal of this withdrawal tendency is restoration. In distancing oneself from the world, one rests, preserves energy, and restores oneself to a neutral position, ready to engage with the world again. This interpretation is also consistent with scientific findings about the evolutionary functions of depression as an emotion, including preventing further actions that might be detrimental to one's health, preventing wasted effort, and conserving energy (Nesse 2000; Beck and Bredemeier 2016).

On the second interpretation, the person might despair about a collection of specific events that she saw from the "24-hour news coverage". If that is the case, the revised argument can explain away this objection. First, as discussed in P2, we should look at emotions through the lens of action tendencies. In this case, despair plays the functional role of prioritizing an action tendency to reduce engagement to achieve the goal of restoration. Specifically, moral despair has caused the agent to avoid hearing or thinking about the humanitarian crises in the world, all of which help achieve the goal of restoring oneself. In distancing oneself from the heartbreaking news, one rests, preserves energy, and restores oneself, until ready to engage with the world again.

Next, notice that the specific objection that C&C make in this example is that "a moral emotion with negative affect fails to motivate ... [the person] to engage in activism" (C&C 2021,

82). As argued in P4, when a judgment that one morally ought to ϕ expresses an emotion, the prescribed action should be a sub-goal that can advance the more abstract goal of the emotion. Accordingly, if the person makes the moral judgment that “I morally ought to engage in activism,” then the actions that fall under the category of activism, such as volunteering and campaigning should all help achieve the goal of the constituting emotion. Notice that these action items all require the agent to engage with the world, which is the opposite of the goal of despair, whose goal is to allow the agent to take a break and get recharged. In other words, the moral judgment “I morally ought to engage in activism” cannot be constituted by a disposition to feel morally despaired.

So, what is the emotion that this moral judgment consists of? In practice, it might vary but the key point is that the goal of the emotion must be in line with the action goal, which is to engage in activism. Some candidate emotions could be compassion or love for humanity or anger. Love for humanity disposes one to care for and help those in need in the moral community. Anger disposes the agent to intervene with the obstacles, which in this case are the social issues that they see on the news. Hence, the agent engages in activism to achieve the goal of removing the obstruction. The reason why this person temporarily does not end up engaging in activism is not because despair fails to motivate them. Rather, it is because their despair is stronger than the emotion(s) that actually constitute the moral judgment that they morally ought to engage in activism. As a result, they are more disposed to withdraw from the world than to step out of their room to fight for their cause.

6 CONCLUSION

In conclusion, C&C's objections to emotion-based accounts of internalism are misplaced. Emotion-based internalists do not argue that affect, as understood by C&C as pleasant or unpleasant mental states, is the source of the motivational power of emotions. Moreover, philosophical scholarship on emotions provides ample theoretical resources for emotion-based arguments for internalism to draw from in order to argue that emotions necessarily motivate actions. The moment we start to understand emotion through the lens of action tendencies and understand the motivational force of emotions as causing behaviors that fulfill certain goals of the emotion, it becomes clear that emotions directed at specific intentional objects necessarily motivate actions. Moreover, when we make a first-personal action-guiding moral judgment, we utilize circumstantial information and practice reason to specify what action we should take, which is an instantiation of the goal of the underlying emotion. As the prescribed action necessarily fulfills the goal of the emotion, we are necessarily disposed to carry out the action.

Thus, to better understand how moral judgments motivate people to act, it is not enough to focus on moral reasoning alone, as the literature on internalism usually does. We should look more into the underlying emotions that give rise to our moral judgments.

Furthermore, connecting the debate on moral motivational internalism with the literature on emotion could be a fruitful research direction in other ways. For instance, it has the potential of enlightening us about how to best carry out moral education in society. If emotion-based approaches to internalism have some truth in them, then the society should focus on cultivating among people appropriate emotional responses to various situations of moral importance. For example, in a misogynist culture, showing compassion or empathy towards women might be frowned upon, which leads to the active suppression of such emotions. As part of the solution to

rectify a misogynist environment, we should collectively recognize the appropriateness and the importance of compassion and empathy towards those who are suppressed.

Another implication of the emotions-based approach is that it helps us identify cases where emotions are unlikely to help prevent wrongdoings. Hence, we should rely on alternative psychological mechanisms to deter crimes. In white-collar crimes, for example, because criminals do not directly confront the victims of their crime, they are less likely to have emotional responses when conducting crimes. A corporate employee who might feel disgusted or outraged by violent crimes, for example, might not feel the same way when they engage in insider trading. In so far as white-collar criminals are less likely to have the appropriate moral emotions, there is less motivational force in refraining from committing the crimes. Policymakers, as well as managers of firms, should pay more effort into building other types of incentive mechanisms to help prevent white-collar crimes. For example, there should be more economic incentives that encourage the reporting of such internal misconduct or crimes. In all, as the motivational power of emotions cedes, it is important that other motivational forces kick in to fulfill the motivational gap such that crimes can be better prevented.

BIBLIOGRAPHY

- Beck, Aaron T., and Keith Bredemeier. “A Unified Model of Depression: Integrating Clinical, Cognitive, Biological, and Evolutionary Perspectives.” *Clinical Psychological Science* 4, no. 4 (July 2016): 596–619. <https://doi.org/10.1177/2167702616628523>.
- Corns, Jennifer. “Unpleasantness, Motivational *Oomph*, and Painfulness: *Unpleasantness, Motivational Oomph*, and Painfulness.” *Mind & Language* 29, no. 2 (March 2014): 238–54. <https://doi.org/10.1111/mila.12048>.
- Corns, Jennifer, and Robert Cowan. “Moral Motivation and the Affective Appeal.” *Philosophical Studies* 178, no. 1 (2021): 71–94. <https://doi.org/10.1007/s11098-020-01421-2>.
- Dancy, Jonathan. “Intuition and Emotion.” *Ethics* 124, no. 4 (July 2014): 787–812. <https://doi.org/10.1086/675879>.
- Deonna, Julien A., and Fabrice Teroni. “Emotions as Attitudes: Emotions as Attitudes.” *Dialectica* 69, no. 3 (September 2015): 293–311. <https://doi.org/10.1111/1746-8361.12116>.
- Eggers, Daniel. “Unconditional Motivational Internalism and Hume’s Lesson.” In *Motivational Internalism*, edited by Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund, 85–107. Oxford University Press, 2015. <https://doi.org/10.1093/acprof:oso/9780199367955.003.0005>.
- Frijda, Nico H. *The Emotions*. The Emotions. Paris, France: Editions de la Maison des Sciences de l’Homme, 1986.
- Kauppinen, Antti. “A Humean Theory of Moral Intuition.” *Canadian Journal of Philosophy* 43, no. 3 (2013): 360–81. <https://doi.org/10.1080/00455091.2013.857136>.
- Kauppinen, Antti. “Moral Sentimentalism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2022. Metaphysics Research Lab, Stanford University, 2022. <https://plato.stanford.edu/archives/spr2022/entries/moral-sentimentalism/>.
- Lazarus, Richard S. *Emotion and Adaptation*. Emotion and Adaptation. New York, NY, US: Oxford University Press, 1991.
- Mitchell, Jonathan. “The Bodily-Attitudinal Theory of Emotion.” *Philosophical Studies* 178, no. 8 (August 2021): 2635–63. <https://doi.org/10.1007/s11098-020-01567-z>.

- Nesse, Randolph M. "Is Depression an Adaptation?" *Archives of General Psychiatry* 57, no. 1 (January 1, 2000): 14–20. <https://doi.org/10.1001/archpsyc.57.1.14>.
- Prinz, Jesse. "An Empirical Case for Motivational Internalism." In *Motivational Internalism*, edited by Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund, 61–84. Oxford University Press, 2015. <https://doi.org/10.1093/acprof:oso/9780199367955.003.0004>.
- Prinz, Jesse J. *Gut Reactions: A Perceptual Theory of Emotion*. Philosophy of Mind Series. Oxford: Oxford University Press, 2004.
- Prinz, Jesse. "The Emotional Basis of Moral Judgments." *Philosophical Explorations* 9, no. 1 (2006): 29–43. <https://doi.org/10.1080/13869790500492466>
- Rosati, Connie S. "Moral Motivation." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University, 2016. <https://plato.stanford.edu/archives/win2016/entries/moral-motivation/>.
- Pool, Eva, Vanessa Sennwald, Sylvain Delplanque, Tobias Brosch, and David Sander. "Measuring Wanting and Liking from Animals to Humans: A Systematic Review." *Neuroscience & Biobehavioral Reviews* 63 (April 2016): 124–42. <https://doi.org/10.1016/j.neubiorev.2016.01.006>.
- Robinson, T. E., and Kent C. Berridge. "The Neural Basis of Drug Craving: An Incentive-Sensitization Theory of Addiction." *Brain Research Reviews* 18, no. 3 (1993): 247–291. [https://doi.org/10.1016/0165-0173\(93\)90013-P](https://doi.org/10.1016/0165-0173(93)90013-P).
- Robinson, M. J. F., A. M. Fischer, A. Ahuja, E. N. Lesser, and H. Maniates. "Roles of 'Wanting' and 'Liking' in Motivating Behavior: Gambling, Food, and Drug Addictions." In *Behavioral Neuroscience of Motivation*, edited by Eleanor H. Simpson and Peter D. Balsam, 27:105–36. Current Topics in Behavioral Neurosciences. Cham: Springer International Publishing, 2015. https://doi.org/10.1007/7854_2015_387.
- Rosati, Connie S. "Moral Motivation." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University, 2016. <https://plato.stanford.edu/archives/win2016/entries/moral-motivation/>.
- Scarantino, Andrea. "The Motivational Theory of Emotions." In *Moral Psychology and Human Agency*, edited by Justin D'Arms and Daniel Jacobson, 156–85. Oxford University Press, 2014. <https://doi.org/10.1093/acprof:oso/9780198717812.003.0008>.

- Scarantino, Andrea, and Ronald de Sousa. "Emotion." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2021. Metaphysics Research Lab, Stanford University, 2021. <https://plato.stanford.edu/archives/sum2021/entries/emotion/>.
- Shargel, Daniel and Jesse Prinz, 2018, "An Enactivist Theory of Emotional Content", in Hichem Naar and Fabrice Teroni (eds.), *The Ontology of Emotions*, Cambridge: Cambridge University Press, 110–129. doi:10.1017/9781316275221.007
- Strawson, Peter. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1962): 187–211.
- Teroni, Fabrice. "Valence, Bodily (Dis)Pleasures and Emotions." In *Philosophy of Suffering*, edited by Michael S. Brady, David Bain, and Jennifer Corns, 103–22. Routledge, 2019.
- Yip, Brandon. "Emotions as Modulators of Desire." *Philosophical Studies* 179, no. 3 (March 2022): 855–78. <https://doi.org/10.1007/s11098-021-01697-y>.