

## **The nature of emotions**

### **Comments on Martha Nussbaum's *Upheavals of Thought: The Intelligence of Emotions* (preprint)**

Joe Lau

Department of Philosophy, HKU

This paper is a discussion of Professor Nussbaum's theory of emotions in *Upheavals of Thought*.<sup>1</sup> This large volume testifies to Professor Nussbaum's impressively wide-ranging interests. It deals with a variety of topics including the philosophy and psychology of emotions, the nature of compassion and its role in social institutions, and the idea and history of love in literature and music. Here I shall focus only on issues surrounding the ontology of emotions.

Nussbaum's theory of the emotions draws heavily on the Stoic account. In her theory, emotions are a kind of value judgment or thought. This is in stark contrast to the well-known proposal from William James, who took emotions to be bodily feelings. There are various motivations for taking emotions as judgments. One main reason is that emotions are intentional mental states. They are always *about* something, directed at particular objects or state of affairs. For example, fear seems to involve the anticipation of danger. To grieve for the passing of a loved one involves the thought that someone dear to us is now gone. In *Upheavals of Thought* and also in her Hochelaga Lecture, Nussbaum analyzed compassion as a set of judgments, including for example the judgment that someone is experiencing serious suffering, and that the person in question does not deserve the suffering.

One might argue about the necessity and sufficiency of these individual definitions. For example, with respect to compassion, Nussbaum assumes that we can only have compassion for the innocent, or those whose suffering go beyond what they deserve. But is this really the case? In the Buddhist tradition, a Bodhisattva is an enlightened being who is supposed to display compassion towards all living creatures, including those who choose to indulge in their desires and deserve to suffer as a result of their own free choices. Perhaps it might be replied that "mercy" is a better word for this type of emotion that does not require non-desert. However, being merciful does not seem to require kindness and concern, whereas compassion does. In this respect the Buddhist conception appears to be closer to compassion.

In any case, the aim of this paper is to discuss the general theory, rather than the definitions for the individual emotions. In philosophical discussion, theories that define emotions in terms of judgments or thoughts are often called "cognitivist". The standard objection to cognitivism is that it cannot

---

<sup>1</sup> Nussbaum (2001) *Upheavals of Thought: The Intelligence of Emotions* Cambridge University Press.

account for the subjective feelings of emotions. For example, how can we explain the overwhelming nature of grief and its heartbreaking feelings, if it is nothing more than a value judgment that is subject to calculation and dispassionate reasoning?

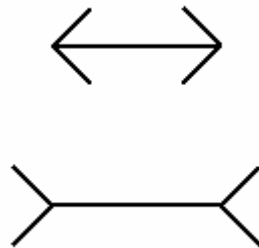
Professor Nussbaum offers at least two replies. First, the experience of an emotion must involve a value judgment where the relevant value is above a certain threshold of importance. To use her example, we do not grieve over the disappearance of a trivial toothbrush. Second, emotions are supposed to involve *Aristotelian* value judgments. When a person has a positive emotion directed at an object, it requires not just the judgment that the object has value. There must also be a *eudaimonistic* judgment to the effect that the object is valuable *to his or her well-being*. When we feel compassion for victims of natural disasters, we are not just acknowledging the loss of values in the lives of the innocent. We are also acknowledging at the same time that these values matter to us, and it is this eudaimonistic element that is supposed to explain the potentially overwhelming power and feelings of emotions.

I am sympathetic to the idea that emotions have a cognitive component, but I also have a number of worries about a pure cognitivist theory of emotions. Some of these worries are general, and some are specific to Professor Nussbaum's theory. Let me start with a general worry, concerning emotional reactions to fictional work such as literature and film. When I watch a movie, I might experience fear and sadness. But I know very well that the characters and events in the movie are not real. It seems wrong to say in such cases that I am making the judgment that someone is in danger, or is suffering in a way that matters to me.

Nussbaum proposes three mechanisms to explain these emotions. First, the spectator might be identifying with the fictional character. Or she might be adopting the perspective of the implied author of the work. Or perhaps her emotion is not so much a reaction to the events depicted in the work, but a response to the potential real-life possibilities suggested by such fictional events. I think these three mechanisms can indeed explain a lot of cases, but I am less certain that they provide a complete account of the phenomenology of emotions in literature and fiction. It seems quite possible to have a situation where I experience fear while watching a movie, but the object of my fear is not considered to be dangerous by the characters or the narrative of the movie. Nussbaum might propose that this is precisely a case where the emotion is actually a judgment about possible similar threats in real life suggested by the movie. I am not convinced that this explanation is always applicable. If I am completely preoccupied by the movie and immobilized by fear, it does not seem plausible from the subjective perspective that I must be making a calculated judgment about the likelihood of a similar encounter in real life. It might even be the case that the fearful situation is one that I believe would never happen to me. But still the fear can persist. If these observations are correct, then genuine judgments are probably not necessary for emotions.

This brings up the question of what a judgment is supposed to be. Let us consider the Stoic account of judgment that Nussbaum uses to introduce her theory. According to the Stoics, a judgment is a two-stage process involving assent to an appearance. It seems to me right now that I am typing in front of a computer. If I accept this appearance as the way things really are, then I come to form a judgment about my present situation. On the other hand, if I am in my skeptical philosophical mood, taking seriously the possibility that I am in a dream, then I might refuse to endorse the appearance and no judgment is formed. I would argue that the emotional reactions to fictional work are more like the Stoic's appearances rather than the subsequent judgments. A gripping horror movie might cause a disposition to judge that there is danger, but it is an appearance that I deny and I don't really judge that danger is present.

To use a perceptual analogy, consider the Müller-Lyer illusion:



When we first encounter this illusion, we might be surprised to learn that the two straight lines are actually of the same length, but a simple measurement soon proves that they are. At this stage, surely we do not have conflicting judgments about the lengths of the lines. Once I have measured the lines, I come to judge that they are identical in length, and this is the only relevant judgment I have. Of course, there remains a disposition or tendency to form the contrary judgment. The lines still *seem* to be of different lengths even though I know they are not. But having a disposition to form a judgment is not the same as having a judgment. This disposition corresponds to the Stoic's appearance. Although I can successfully resist the power of the appearance, I cannot get rid of it.

The perceptual analogy is relevant to the analysis of situations where emotions persist despite contrary judgments. If I am afraid of heights, I might be paralyzed by fear when I stand on the balcony of a tall building, even though I know I am perfectly safe. I think Nussbaum would say this is a case of conflicting or oscillating judgments. This might well be the case. But intuitively, it is also possible that I am not being in any sense irrational, that I do not waver in my judgment that the balcony is perfectly safe. What I am battling with is rather a psychological propensity to believe otherwise, which, to my credit, fails to convert into a genuine judgment. To settle this dispute, we need to specify more clearly what a judgment is. One option is to take a judgment to be an "all-things-considered" occurrent belief to which we would normally give verbal assent, in which case

there are no conflicting judgments in the situation just described. Or one might adopt a more liberal perspective, including as judgments related pro-attitudes that generate propensities to action, even though we might reject them on reflection. In that case, recalcitrant emotions do indeed involve conflicting judgments, but then emotions should be compared with the Stoic's appearances rather than the subsequent affirmations that Nussbaum favors.

Let me now turn to a more specific worry about the role of *eudaimonistic* judgments in animal emotions. Against the Stoics, Professor Nussbaum argues convincingly that many animals can experience emotions. But the problem is that these animals might not be capable of eudaimonistic judgments, and the contents of eudaimonistic judgments make necessary reference to the concept of the self. Although there is some empirical evidence that dolphins and orangutans have a concept of the self, this might not be true of other animals that can experience emotions. An animal might be able to judge "here is food", but it might be unable to form the more sophisticated judgment that "here is food *that is good for me*." If I understand Nussbaum's proposal correctly, for animals to have emotions, they must be capable of eudaimonistic *appraisals* rather than judgments. The contents of these appraisals need not be linguistic propositions, but they still require aspect-seeing ("seeing X as Y") - the ability to see certain objects or situations *as* salient or important for the well-being of the creature.

I have two worries about this proposal. First, aspect-seeing is still conceptualized seeing. Seeing a cloud in the sky that happens to look like a donkey is not the same as seeing the cloud *as* a donkey. The latter requires the concept of a donkey, since to see X as Y is to represent X as having property Y rather than some other property, and it is not clear how this can happen without the subject possessing a concept of Y. If this is correct, eudaimonistic appraisals in animals still require a concept of the self and a concept of well-being. These might be rather minimal concepts in the case of animals, but I still wonder whether they are necessary for animal emotions. Having a system of beliefs and desires that feel something to the creature might be sufficient for some emotions, but it would not be sufficient for any conception of well-being or self. In any case, the other issue to consider is that we end up with a disjunctive theory of emotions, one for human beings and one for animals. But if full eudaimonistic judgments are not necessary for animal emotions, we might begin to wonder why they are necessary for humans. It might be true that emotions are value judgments, and furthermore, it might also be true that the values involved in these judgments must be values that have a personal significance. But it does not follow that emotions must include the specific judgment that *these are values that are important to me*. Giving up this stronger requirement is compatible with thinking that value judgments in emotions have a personal dimension, and it also allows a unitary theory of emotions that applies to both humans and other animals.

I have raised a general worry about the case of fiction, and a more specific worry about the eudaimonistic requirement. But for many philosophers, the most serious objection to cognitivism is that it cannot account for the subjective feelings of emotions, what it is like for an agent to consciously experience emotions. This objection applies to Nussbaum's theory as well, and I would like to examine this objection in more details and hopefully also offer some constructive comments.

Let us focus on the case of the attenuation of grief over time, a topic that Nussbaum examines in depth. When someone we love passes away, the emotional impact might gradually diminish over time. The initial sorrow and paralyzing misery might give way to contemplative longing and a more subdued sense of emptiness. What is undeniable is that the emotion of grief can remain even as its intensity changes. Yet for all we know, our evaluative judgments concerning the value of the person we mourn might not have changed. In fact, as our grief subsides, our appreciation of that person might actually increase as we look back objectively and come to a deeper understanding of that person's contribution. This implies that grief cannot be an evaluative judgment since the way its intensity changes does not correspond systematically to the value we assign in our judgments.

Nussbaum discusses this objection at some length, and I cannot do justice to her response within the limited space of this discussion. The key to her response lies in the observation that a person who is valuable to us is often a person who is closely involved in our lives and associated with activities that we treasure. She says it is crucial to recognize that the eudaimonistic elements of the relevant judgments have changed. Before the person passed away, I judge "this person *is* an important element in my flourishing". But now the judgment becomes "the person who died *was* a central part of my life." According to Nussbaum, this change of tense of the judgment "is a large part of what constitutes the diminution of grief." As reality sinks in and we revise our eudaimonistic judgments about the role of that person in our on-going projects, the intensity of grief decreases.

I am not convinced that these changes can explain the decrease in emotional intensity. Although grief might attenuate over time, its intensity can still fluctuate depending on contingent circumstances which have nothing to do with eudaimonistic assessments. Although a beloved family member might have past away a long time ago, an old photo can still trigger an intense feeling of grief on some occasions, even though all the relevant eudaimonistic judgments are already in the past tense.

Let us consider a similar example involving anger. Suppose I just discover that someone tried to harm me in an underhanded manner a long time ago. Although that person failed to affect my well-being, and it was a long time ago, I might still be very angry when I first learnt of the incident. This anger that I experience might attenuate over time, even on those occasions when I think about the incident again and experience the emotion in an occurrent rather than background manner. But all my evaluative judgments remain the same past-tense judgments as before.

The problem posed by the examples of grief and anger is a serious one: if the values that I assign in my judgments do not correspond systematically to the intensity of my emotions, why should we think that having value judgments are *ever* sufficient for the experience of any emotion?

A typical way of dealing with this line of objection is to incorporate subjective feelings into a hybrid theory of emotions. Perhaps emotions are complex mental states including both bodily sensations and judgments. Robert Solomon, a well-known defender of pure cognitivism, is now inclined to accept the hybrid approach. It is beyond the scope of this paper to discuss the pros and cons of the hybrid approach, but perhaps I can offer a simple if radical way out for the cognitivist.

My tentative proposal for the cognitivist is to bite the bullet and agree that evaluative judgments cannot explain how emotions feel to us. In fact, evaluative judgments are not sufficient for the experience of emotions, and this is because emotions are actually *always* unconscious. However, these unconscious emotions can cause conscious emotional feelings. When we experience an emotion, there are two things that happen: we have an unconscious emotion constituted by one or more evaluative judgments, and this emotion *causes* a set of conscious feelings which are symptomatic but not constitutive of the emotion. In other words, emotions and emotional feelings are causally related, but they are distinct from each other.

Why should a cognitivist think that all emotions are unconscious? The argument has two premises. The first premise is that emotions are identical to evaluative judgments, which is what cognitivism is. The second premise is another seemingly radical claim that I shall defend shortly. The claim is that *all* judgments and thoughts are actually unconscious, whether they are evaluative or not. Putting the two premises together, we get the conclusion that all emotions are unconscious.

It is natural to think that the second premise flies in the face of commonsense and introspection. If we are not conscious of our thoughts, how do we know what we are thinking? Also, many people experience thinking as a conscious internal monologue. For others, thoughts seem to present themselves as fleeting images occurring in our consciousness.

However, my response is that these verbal and perceptual images are the causal effects of our thoughts and are not identical to them.<sup>2</sup> As Wittgenstein reminds us in the *Philosophical Investigations*, when we think of an old man walking up the hill and form a mental image of the situation, the image is not to be confused with the thought because the image might as well be an image of an old man sliding backwards down the hill. Similarly, when thinking takes the form of a

---

<sup>2</sup> This position was defended by Jackendoff in his (1990) *Consciousness and the Computational Mind* Cambridge: MIT Press.

silent monologue, the sequence of speech sounds we hear in our minds can express very different thoughts in the minds of the speakers of other languages. In other words, the speech sounds or images that we experience are indeterminate in meaning in a way that our thoughts are not. In any case, whatever contents these images have, they are distinct from the contents of our judgments. So we do have some independent reason for thinking that all judgments and thoughts are unconscious. If emotions are evaluative judgments, they would have to be unconscious as well.

This is not the place to provide an extended defense of this controversial theory, but I think the cognitivist should welcome such a proposal. First, the thesis that emotions are evaluative judgments can still be maintained. Second, the objection that evaluative judgments cannot explain emotional feelings is now seen to be completely misplaced. By distinguishing between emotions and emotional feelings, the cognitivist can now allow a far more complex relationship between the contents of evaluative judgments and the subjective phenomenology of emotions. For example, Nussbaum can still claim that in normal situations, the intensity of an emotional feeling varies systematically with the amount of eudaimonistic value specified in the relevant judgment. But perhaps perception and memory can also affect the perceived intensity. By taking all these factors into account, we can provide a more natural explanation of the attenuation of grief and anger over time. Of course, the present proposal does imply that our ordinary talk about conscious emotions is strictly speaking mistaken. But we can easily reinterpret “conscious emotions” in ordinary language to mean “unconscious emotions that we are indirectly conscious of (through the emotional feelings they cause).” This still allows the possibility of unconscious emotions at a deeper level, to which we do not even have indirect access because they do not give rise to emotional feelings.

Obviously, there is a lot more that can be said about this proposal. In particular, I have yet to explain the nature of emotional feelings, such as whether these feelings are simply bodily sensations. I would argue that emotional feelings would have to include affective qualities such as the feelings of pleasure and displeasure, and perhaps some other ones as well. But again these conscious affective qualities are not to be confused with the emotions proper. I realize that these brief remarks are no more than a framework for thinking about the emotions, but I hope I have said enough to convince those who are sympathetic to cognitivism that it is worthy of their consideration.