

Do Unconscious Emotions Involve Unconscious Feelings?

Michael Lacewing

The very idea of unconscious emotion has been thought puzzling. But in recent debate about emotions, comparatively little attention has been given explicitly to the question. I survey a number of recent attempts by philosophers to resolve the puzzle and provide some preliminary remarks about their viability. I identify and discuss three families of responses: unconscious emotions involve conscious feelings, unconscious emotions involve no feelings at all, and unconscious emotions involve unconscious feelings. The discussion is exploratory rather than decisive for three reasons. First, the aim is to provide a framework for the debate, and identify a number of key issues for further research. Second, a number of the positions depend for their plausibility upon theoretical commitments that can be made clear, but cannot be evaluated in detail, in a survey article. Third, I believe no fully satisfactory, comprehensive solution has yet been developed.

Keywords: Consciousness; Emotion; Feelings; Unconscious Emotions; Unconscious Feelings

1. Introduction

In the flurry of recent debate about the nature of emotions, comparatively little attention has been given to the question of unconscious emotions. But it is clear that we need the concept in order to make sense of human behavior. It is commonplace for people to later realize what they felt, but were not consciously aware of feeling, at an earlier time; everyday explanations of people's behavior, perhaps particularly in personal relationships, require us to attribute to them emotions of which they are not aware; literature is full of examples and illustrations of characters' ignorance and self-deception regarding what they feel. Yet the philosophers and psychologists who do at least explicitly mention unconscious emotions spend relatively little time developing their position and defending it against objections.

Correspondence to: Michael Lacewing, Heythrop College, University of London, Kensington Square, London W8 5HQ, UK. Email: m.lacewing@heythrop.ac.uk

Perhaps part of the general reticence on this issue has stemmed from the thought that there is very little to discuss, that the concept of an unconscious emotion is impossibly paradoxical; and so however we are to make sense of people's emotional lives, this concept will have no role. Predominant in this line of thought is the notion that "unconscious emotion" involves a commitment to unfelt feeling; and this is taken to be so obviously nonsensical, few philosophers have felt the need to provide an argument against it (Clore, 1994). The very idea of unconscious emotions has been thought puzzling even within what is perhaps its most natural home, psychoanalysis. Freud (1915/1984) remarked, "It is surely the essence of an emotion that we should be aware of it, i.e. that it should become known to consciousness" (p. 179).

This article surveys a number of recent attempts by philosophers to resolve the puzzle and provides some preliminary remarks about their viability. The discussion is exploratory rather than decisive for three reasons. First, the aim is to provide a framework for the debate, and identify a number of key issues for further research. Second, a number of the positions depend for their plausibility upon theoretical commitments that cannot be evaluated in detail in a survey article. In these cases, my aim is the more limited one of situating the accounts within such commitments. The third reason the discussion is not an attempt to reach a decisive conclusion is that I believe no fully satisfactory, comprehensive solution has yet been developed.

We can divide up answers to the title question into three families. The first two reject the idea of unconscious feelings. The first family claims that unconscious emotions involve conscious feelings. This family has two branches. The first branch, of which the views of Aaron Ben-Ze'ev (2000) and Patricia Greenspan (1988) are examples, claims that the conscious feelings are misunderstood (§5). The second branch, exemplified by Peter Goldie (2000), defends a distinction in consciousness, arguing that unconscious emotions are consciously felt, but that the subject remains unaware of the feeling (§6).

The second family of theories denies that unconscious emotions involve feelings at all. It also has two branches, represented by Martha Nussbaum (2001) and Robert Roberts (2003), each of whom provides a different kind of argument for separating emotions from feelings (§§8–9).

The third family defends the idea of unconscious feelings, and argues that they are involved in unconscious emotions, usually on psychoanalytic grounds (§10). This family includes the theories of Sebastian Gardner (1993) and Richard Wollheim (1984).

This division into three families is further complicated by those theories that seek to use physiological or neurophysiological criteria for emotion and feeling rather than purely psychological criteria. Interpretations of the theories of Antonio Damasio (1994, 2000) are possible within each of the families distinguished, depending on how the relation between bodily or neurophysiological change and consciousness is conceived (§§11–12).

There are, of course, other important theorists presenting arguments within each of these families. This survey is not comprehensive in that sense. My main concern is

with the families and their branches as potential solutions to the puzzle of unconscious emotions, rather than with individual differences between theories within each branch. So I shall use just those theories mentioned to exemplify the type of solution offered in each case and the objections it faces.

2. What is an Unconscious Emotion?

2.1. A Definition

It would be helpful to have a definition of ‘unconscious emotion’. Unfortunately, any informative definition will already be committed to a certain position, or at least family, in the spectrum of possibilities. The only point on which all theories agree is that an unconscious emotion is *an emotion that the subject is not aware of in such a way as to be able to avow it directly and noninferentially* (the last three words are intended to rule out a case of inference from one’s behavior). Whether and how the subject is aware of it in any way at all is contentious, and an important aspect of the question we face.

Like all definitions relating to consciousness, there is a puzzle about how the *epistemological* and *phenomenological* aspects of the definition relate. The concept of “feeling” reflects this. Commonsense suggests that the type of awareness involved in feeling an emotion essentially involves phenomenology, i.e., the epistemological access that enables direct and noninferential avowal of an emotion necessarily involves, even if it cannot be reduced to, “feelings.” I shall assume that this account of feeling an emotion is correct and will therefore talk of feeling (verb) an emotion as involving feelings (noun). However, I leave it open, as a matter of dispute, as to whether “feelings” are or involve qualia, or are entirely reducible to intentional content.¹

It will emerge in the course of the argument that I believe that different accounts of why or how an emotion is unconscious are appropriate in different cases. In other words, there are different ways in which an emotion may be unconscious; there is not just *one* description of the subject that is applicable in all cases. A theory of unconscious emotion therefore needs to be able to encompass these different ways in which emotions may be unconscious. This will form an important criticism of the first family of theories, which claim unconscious emotions involve conscious feelings. These theories *rule out the possibility* that unconscious feelings can occur. By contrast, the theories that allow for unconscious feelings do not reject the claim that unconscious emotions may also involve conscious feelings. The latter theories are therefore more ecumenical, which may prove a strength.

2.2. An Example

In the absence of a fuller definition of unconscious emotion, we can at least begin with a case to think about, to sharpen our intuitions against. Freud’s famous case study *Notes on an Obsessional Neurosis* (1909/1979), better known as “the Rat Man,”

provides helpful instances. I choose a psychoanalytic case study at least in part because emotions that are unconscious in the manner exemplified by psychoanalytic case studies are the most difficult to account for. Some theorists may feel that if their views can't account for this type of "psychoanalytically unconscious" emotion, so much the worse for the claim that such emotions exist. However, I find the concept as used in the case study below independently plausible, on the grounds that I cannot see how to make sense of human behavior without it.

The "Rat Man" was a man in his late 20s, training as a lawyer at university, who suffered from obsessive-compulsive disorder. His case name derives from a powerful and recurrent fantasy or fear that precipitated his seeking treatment with Freud. His captain in the army had told him of a punishment used in the East—a pot containing rats is placed on the buttocks of the prisoner, and they bore their way into his anus. When the Rat Man heard this, it flashed through his mind that this was happening (inflicted impersonally) to the two people he loved most, to the woman he was romantically involved with and to his father.

On a day when the Rat Man's lady was leaving after visiting him, his foot knocked a stone, and he felt obliged to move it to the side of the road in case her carriage, which would pass that way, would strike the stone and be overturned. He walked on a bit, but then thought his action was ridiculous, and he felt obliged to go back and put the stone back in the middle of the road.

Now most of us, when we feel we've done something ridiculous, don't feel a compulsion to go back to the scene of action and "undo" it! So, as Freud (1909/1979) says:

We shall not be forming a correct judgement of this second part of the compulsive act if we take it at its face value as having merely been a critical repudiation of a pathological action. The fact that it was accompanied by a sense of compulsion betrays it as having itself been a part of the pathological action, though a part which was determined by a motive contrary to that which produced the first part. (p. 72)

On the basis of a great deal more information than we have time to rehearse, it became clear that the first gesture expressed the Rat Man's conscious love; the "undoing" action expressed his unconscious hatred. Freud goes further: his obsession for protecting his lady was actually a reaction to his unconscious hatred, which threatened his love. His hatred, therefore, was in fact active in his removing the stone as well as in his replacing it (see Wollheim, 1984, ch. 5).

3. Emotions: Episodes and Dispositions

To think about what unconscious emotions might be and whether they involve unconscious feelings, we need first to make an important distinction which is not always noted in the literature. Goldie (2000, pp. 12–16) argues for a distinction between *episodes of emotional experience* and emotions themselves. Many contemporary analyses of emotion are in fact analyses of episodes of emotional experience. Goldie, however, argues that we attribute emotions on the basis of their

place in a narrative. For example, if a man is jealous of a rival—an emotion that may last for years—this does not consist simply in an episode, or even several episodes, of his consciously feeling jealousy, but also in thoughts, other feelings, and bodily states, and dispositions to all these. What makes the emotion jealousy is given by the terms in which we understand him, his life, and his relations to the object of his desire and his rival. Emotions, and the ascription of emotions, find their home in the narrative structure of people's lives.

On this account, emotions are dispositions: “He is proud of his children,” “she is afraid of snakes,” even “he is angry with his boss today.” None of these attributions are reducible to the attribution of a constant, continuing episodic mental state—whether this is thought to be of feeling (Goldie, 2000), thinking (Nussbaum, 2001), or construing (Roberts, 2003). The dispositions attributed are manifest most directly in episodes of felt emotion, and such episodes have a special place in our understanding of what emotion is. But we should not think that such episodes of feeling (thinking, construing, etc.) are all there is to an emotion. This is obvious in being proud of one's children or afraid of snakes, which last a long time. But even in the case of being angry at one's boss for a day, there is an intuitive way of understanding this that is not simply equivalent to a persistent episodic state. The feeling (or episodic thought) may come and go, for instance with natural redirections of attention. Whether I am still angry the next day, for example, is decided (in part) by whether I am disposed to feel angry once again. Whether the disposition begins and ends with the episode of feeling is a *discovery* we need to make; and so even if it does, the two are logically distinct and we can understand the feeling as the manifestation of the disposition.

Emotions also have histories and can undergo change. The emotions we have now, what and how we love, hate, fear, etc. involve accretions, past objects of emotion that color and shape our responses to present objects. Philosophers (e.g., Rorty, 1980) who have built this idea of history into their theory of emotion tend to analyze emotions in terms of the subject's evaluative understanding of the world. Wollheim (1999, 2003) argues that emotions are evaluative attitudes, or orientations, towards the world or objects within it. This is easily combined with Goldie's account: how someone understands the world and what is of importance in it has a history and can be given a narrative form.²

This is the understanding of emotion I shall assume in this paper, and not seek to defend further. The Rat Man's unconscious hatred is a disposition; the episode of his removing and replacing the stone was an episode of unconscious hatred—whether or how it was “felt” is our question. Assuming this analysis does not disadvantage accounts that have analyzed emotions as episodes of emotional experience. These accounts allow that we have dispositions to emotional experience, that these dispositions are closely related to and interact with dispositions to certain types of thought and behavior. Adherents to these accounts can, therefore, take what is said below about emotions as applying to such dispositions. I will make clear when a theorist's analysis of emotion is, in Goldie's terms, an analysis of an episode of emotional experience.

4. Defining the Puzzle

With the distinction between emotion and episodes of emotional experience in place, and understanding emotion as an evaluative orientation, it may seem there is no puzzle as to how emotions could be unconscious. For there is no reason to think that we are aware of our evaluative orientations to the world; and plenty of reasons to believe that not only are we often unaware of such orientations, but we are motivated to be so. Our evaluations of others, ourselves and the situations we face can be painful, either for what they reveal about the situation or about ourselves. As a result, we avoid recognizing them.

But although this is correct, it overlooks the rest of Goldie's analysis, viz., that these orientations are dispositions to, among other things, episodes of emotional experience, which, until we have an argument to believe otherwise, we should take to mean episodes of *feeling* that emotion. If the emotion—the evaluative orientation—is to remain unconscious, then any episodes of feeling that emotion cannot become (fully) conscious; the subject cannot be fully cognizant of the feeling and what it is a feeling of, or she would thereby become fully cognizant of the emotion itself, and be able to avow it.

This leads to the three solutions presented in the introduction: unconscious emotions are felt consciously, but the subject is not fully cognizant of them in some way (family 1); or unconscious emotions involve unconscious feelings (family 3); or they do not involve feelings at all (family 2).

On the type of evidence appealed to in the first paragraph, and cases such as the Rat Man, we should not adopt a *purely* dispositional account of unconscious emotion. The same considerations that lead us to posit unconscious emotions in the first place also incline us to attribute *episodes* of emotion of which the subject is unaware (this is inherent in Freud's idea of the "dynamic unconscious"). Since the term 'experience' has connections to consciousness, let us say these are episodes of unconscious emotional "activity." Unconscious emotions are dispositions to such, among other manifestations. How we should understand these episodes is the question.

It is obvious that the Rat Man was undergoing an episode of emotional feeling while moving the stone—his love, his sense of compulsion, his sense of his action being ridiculous were all felt. But what is missing from his report is his feeling hatred; this he denied. Yet his hatred was "active" in his moving the stone; it was not just a disposition, but a disposition *being manifest at that moment*. His unconscious emotion of hatred was expressed in action and had effects upon his conscious mental states. In his interpretation, Freud relates the feelings of protective love, compulsion, and the sense of ridiculousness to the unconscious hatred. Should we say that the episode of unconscious hatred occurred without any feeling of hatred, e.g., as an episodic thought without feeling? Or did the Rat Man feel his hatred? If so, did he feel it consciously or unconsciously?

5. Unconscious Emotions Involve Conscious Feelings 1

5.1. Theory

I turn first to those theories that maintain that unconscious emotions are felt consciously. (I interpret the idea of “feeling” psychologically, except in §§11–12, where physiological and neurological theories of feeling will be discussed.) The first theory of this kind I shall call *conscious feelings 1* (CF1).

There is no inconsistency in maintaining that unconscious emotions manifest themselves in conscious feelings. What makes it appear so is the additional assumption that conscious feelings would *reveal* the unconscious emotion, rendering it conscious. But although it is perhaps normal for feelings to reveal the existence and nature of the emotions they manifest, it is obviously true that this is not universally so.

Because episodes of emotional feeling are manifestations of emotion, in the normal case the occurrence of feeling informs us about the emotion—e.g., which emotion we feel, that an old emotion we thought had died is still extant, that a new one has formed, or that a familiar one is present. However, feelings may be confused, unclear, and in need of interpretation. Just to experience the feeling cannot *by itself* inform us of the nature of the emotion. As Wollheim (1999, p. 10, 2003, p. 22) argues, to understand feelings, we must *relate* them to the emotions they manifest. We cannot understand that what we feel is anger unless we come to understand that we are angry. It is only through associating the feeling to the disposition that we are able to fully recognize the feeling for what it is. If we are unable to relate the feeling to the emotion correctly, the feeling remains obscure and the emotion unknown.³

This, then, allows for the possibility that a subject does not understand what he feels. The feeling may be identified as “a feeling” without being identified as the type of feeling it is. It may simply be not understood, or it may be misidentified or misunderstood as manifesting an emotion it does not, in fact, manifest. (As misidentification is a type of misunderstanding, I shall speak just of the latter from now on.) The subject thereby misunderstands the emotion he feels. We should reject Descartes’ (1650/1989) claim that passions “are so close and internal to our soul that it is impossible it should feel them without their truly being as it feels them” (Article 26). It is possible, therefore, for there to be an episode of conscious feeling that manifests an unconscious emotion without the emotion ceasing to be unconscious. It remains unconscious as a result of the subject’s not correctly understanding his feeling; he is unable to associate the feeling to the evaluative orientation on the world that he in fact has, but is unaware of.

We may add that this lack of self-understanding may be a result of psychological defense, particularly in the sorts of cases in which psychoanalysis is most interested. That is to say, the misunderstanding is *motivated* by the wish to avoid mental pain (anxiety, distress, guilt, shame and so on). Because to recognize the emotion would cause pain, it is kept unconscious (on this theory, by misunderstanding the feelings that manifest it). On this account, the Rat Man is conscious of the feelings that

manifest his hatred during the episode described above, but does not understand his feelings in relation to hatred. His hatred therefore remains unconscious.

5.2. Evaluation

As noted at the outset, the view that feelings must be conscious is very widespread amongst philosophers and psychologists. As a result, CF1 is perhaps the default position in the literature. It is therefore worth spending some time developing the debate between its defenders and its detractors.⁴

This account is undoubtedly true in some, perhaps many, cases of what we may call “unconscious emotion.” It is intuitively plausible that to think that while feeling an emotion, we can mistake its object, its significance, or even confuse it with another type of emotion that can feel similar. Self-deception often appears to work on this basis.

But the theory is unsatisfactory as a general account, i.e., it is a correct description of how and why some emotions are unconscious, but it cannot account for all unconscious emotions.

5.2.1. The problem of feelings

A first objection is that its explanation of the case of the Rat Man is implausible, because it requires the Rat Man to *radically* misunderstand his feelings. We are asked to suppose that his *feeling* of hatred can still meaningfully be described as a feeling of *hatred*, and all that changes is the understanding of it. But can the Rat Man—or anyone—mistake the feeling of hatred for, say, a feeling of ridiculousness and compulsion, or even solicitous love? Is it not more plausible that something apart from straightforward misunderstanding, or lack of understanding, is occurring?

A possible reply is based on the claim, which both Greenspan (1988, pp. 4–5) and Ben-Ze’ev (2000, pp. 49–50, 63–66) defend, that emotional feelings on their own carry very little information. They are not “rich” in the intentional content of the emotion, but separate from it, occurring along simple dimensions of pleasure and pain. It is wrong, then, to say that the Rat Man “feels his hatred.” Rather, he is conscious of the feeling “component” associated with his hatred. But his feeling, taken in isolation, is not a feeling “of hatred,” as though hatred could be read off from the feeling. It is perfectly possible, therefore, that feelings associated with hatred, taken in isolation from the emotion, could be misunderstood for the sorts of feelings the Rat Man describes and avows. Feelings only *seem* laden with greater intentional content, if they do at all, once we have linked the feeling to its associated emotion.

However, we may object that making feelings *so* lacking in content empties the claim that the unconscious emotion is felt of meaning. It would be better to say that some generic feeling of discomfort occurs: If feeling is so lacking in content, what distinguishes the claim that the Rat Man feels hatred from the Rat Man feels anxious?

(Perhaps we should say the Rat Man feels anxious *instead of* feeling hatred, i.e., given that we can attribute hatred to him on the basis of the narrative of his life,

we would expect him to feel hatred. But at the moments when we expect that feeling, he reports anxiety instead. This is an interesting response, but to say this is distinct from saying that he feels, but misunderstands, his feeling of hatred. The hatred does not remain unconscious as a result of misunderstood feelings, but because the feelings associated with its episodes of activity are not feelings of hatred, but of anxiety. If this suggestion is right, there is clearly a complex story to be told about the relation between the unconscious hatred, the “missing” feelings of hatred, and the conscious feelings that appear in their stead.)

Second, leaving aside the issue of how to interpret the Rat Man, the defense crucially depends on the contentious claim that feelings have virtually no intentional content. The suggested separation of feeling from content applies not just to unconscious emotion, but is a general account of the nature of all emotion. Several theorists have provided different and independent arguments that feelings are more richly intentional than this theory allows (Goldie, 2000, ch. 3; Nussbaum, 2001, ch. 1, §§4, 6; Wollheim, 1999, ch. 2, §11, 2003). It is correct, they argue, to talk of “feelings of hatred” rather than “the feeling component associated with hatred.” What makes hatred hatred is at least partially manifest in the feeling itself. If this were not so, we might wonder how, in the normal run of things, we are able to identify our emotions from our feelings with such ease. If feelings were so “thin,” much more explicit thought would be necessary to “work out” what emotions we feel. This point does not entail the implausible claim that we can simply “read off” our emotions from our feelings. There is more to go on than “component” theorists think, but feelings still require interpretation. We don’t have space to review this debate here (though see §§8.1 and 10.1 below for points from Nussbaum and Wollheim).

This first objection, then, takes the form of a dilemma: the “thinner” feelings are, the more plausible the claim that we can be conscious of the feeling “component” of an unconscious emotion, but the less plausible the theory of feelings. However, making feelings “thicker” makes it more and more necessary to say that conscious feelings must be radically misunderstood for the associated emotion to remain unconscious.

5.2.2. Defense and the problem of motivation

A second objection, this time from the defenders of unconscious feelings, turns our attention to those cases in which misunderstanding the feeling is motivated in a way psychoanalysis describes, viz., the misunderstanding is the result of a defense mechanism.⁵ Psychoanalysis notes that one reason why emotions are kept unconscious, e.g., as in the case of the Rat Man, is that they are painful. In response to the painful emotion, a psychological defense operates which attempts to reduce or eliminate the pain. Since to feel and understand the emotion consciously would be painful, the defense keeps the emotion unconscious.

According to CF1, feelings are conscious, and so any pain that an unconscious emotion causes would be conscious. Hence, defense can only work if it prevents pain through a lack of understanding of the conscious feelings that manifest the emotion.

In other words, this lack of understanding is caused by an *anticipation* of the pain that true understanding would bring.

The objection that defenders of unconscious feelings bring is that this cannot be an adequate account of defense. Instead, they claim, we must suppose that the emotion *is painful*, and psychological defense operates to prevent this pain from becoming conscious. If they are right, there is unconscious pain, and since pain is a type of feeling, there can be unconscious feelings. Unconscious pain cannot be understood as pain that is felt consciously but misunderstood. Simply *misunderstanding* such pain is insufficient to keep it unconscious—misunderstood pain is still painful, and so CF1 can't account for the pain that motivates our misunderstanding.

I shall return to this argument when looking at unconscious feelings in §9.

5.2.3. *The problem of the absent feeling*

A third and final objection is this: what can this theory say about cases in which the subject reports *no* particular feeling at a time when we would expect and want to attribute an episode of emotional experience? As this forms an objection to the next theory, *conscious feelings 2* (CF2), as well, I shall discuss this in §6.2.2.

6. Unconscious Emotions Involve Conscious Feelings 2

6.1. *Theory*

To develop the idea that being mistaken about one's feelings is not all that there is at issue, we may appeal to a more sophisticated understanding of consciousness. Philosophers and psychologists commonly draw a distinction between two levels of consciousness, usually understood in terms of points on a continuum without a sharp boundary between them. There are different versions of and names for the distinction, but the central idea is relatively clear. I shall adopt Goldie's (2000, pp. 63–70) terms. Goldie distinguishes between “reflective” and “unreflective” consciousness, the former being consciousness of our thoughts and feelings about the world, the latter being consciousness of the world. Not only is *having* an emotion fundamentally a matter of being engaged with the world, he argues, so is *feeling* an emotion. Feeling focuses on its object, not on itself. We may therefore not identify ourselves as having the feelings towards the world we do until or unless we become reflectively conscious of them.

We may doubt that we really have *feelings* when unreflectively engaged with the world. We may grant that we can have thoughts—intentional content—directed onto the world without reflective consciousness; but could such unreflective states also have *phenomenology*? Goldie argues that they do by noting that, when asked to reflect on how we feel, we may be able to say what it is we feel, even though prior to that point, we had been “unaware” of our feelings. This may apply in the moment or it may apply across long stretches of time, as we reflect back on previous episodes of our lives. As I recall the episode, the feeling arises, and I come to realize that the feeling is one I had at the time.

The idea of feelings that are conscious but of which we are unaware can be defended by analogy with perceptual states. Two famous cases in the philosophical literature are that of suddenly realizing that there has been a pneumatic drill operating in the distance for some time, and that we heard it, but had not noticed it (Block, 1997); and that of a long-distance lorry driver who “comes to” after a period of absent-minded distraction and realizes that though he must have seen the road in order to drive safely, he has no recollection at all of doing so (Armstrong, 1968). It can be argued that the sound and the sights were conscious, but the subject was not conscious of the perceptions. The perceptions were conscious not just in the sense of being directed onto the world, but also—being perceptions—in having a phenomenology (e.g., Dretske, 1993).

Goldie’s analysis argues that there are episodes of emotional feeling we undergo without being “reflectively conscious” of them, i.e., we are not aware that we are undergoing such feeling even though there is a legitimate sense in which the feelings are conscious.

This provides an explanation for how it is that unconscious emotions may manifest themselves in episodes of emotional feeling without themselves becoming conscious (while not needing to insist that feelings have little or no intentional content). If the feelings that manifest the emotion are only unreflectively conscious, the subject will not become reflectively conscious of the emotion, and so remain unaware that she has the emotion. Furthermore, the unreflective feelings may be masked by other feelings of which she is reflectively conscious, and this masking may be motivated. Psychological defense, then, may involve the exaggeration of or focused attention on some feelings at the expense of others, or even the creation of factitious feelings—feelings that are not manifestations of the real evaluative orientations the subject takes to the world; all of which serves to obscure from reflective consciousness the feelings that manifest the unconscious emotion. The Rat Man, then, does feel his hatred, but not in reflective consciousness, i.e., he is not conscious *of* feeling hatred. Furthermore, his feeling of hatred is masked by the feelings he is aware of, including his anxiety and his rather solicitous loving concern. (And if we wish, we may say this explains how it makes sense to say the Rat Man feels these feelings “instead of” hatred.)

6.2. *Evaluation*

This model is an apt description of many instances of unconscious emotion, and improves on the previous theory. But it, too, does not cover the whole field, particularly in psychoanalytic cases.

6.2.1. *The disanalogy with perception*

In the examples from perception, if asked to direct her attention to what she heard or saw, the subject could have become reflectively conscious of what was unreflectively conscious. And Goldie uses this fact to argue for unreflectively conscious feelings. It is this argument that establishes their close link with consciousness and protects their

integrity as *feelings*, i.e., having phenomenology. But this prevents the account from providing a satisfactory analysis of unconscious emotion generally, for in the case of the Rat Man, and many others in the psychoanalytic literature, the subject *cannot* simply turn his attention to the feelings he has in unreflective consciousness. The Rat Man, if asked, when moving the stone or afterwards, to direct his attention toward his feelings, would deny that he felt hatred or anything like it.

To save the theory, we may appeal to the fact that the subject is motivated not to bring such feelings into reflective consciousness and so may not be able to do so; and that other, reflectively conscious feelings mask the unreflectively conscious ones. The Rat Man, for instance, does report other feelings that we could plausibly argue stand in the way of his bringing his feelings of hatred to reflective awareness. In the examples from perception, neither of these two facts holds. But now we lack the very source of evidence for unreflectively conscious feelings that Goldie uses to introduce them, and the analogy with perception falters. Without the kind of retrospective confirmation described above, with the subject calling into reflective consciousness their earlier experience of feeling those emotions, why should we think that the unconscious emotion manifests itself in conscious feelings at all?

6.2.2. *The problem of absent feeling*

The objection is sharpened if we consider psychoanalytic cases in which the subject reports no particular feelings *at the time of the emotional episode*, rather than retrospectively. This eliminates the possibility that the lack of confirmation is due to an error of memory. The psychoanalyst Joyce McDougall (1986, ch. 7) describes patients she calls “normopathic.” These patients appear to have very little inner, emotional life. They have a tendency to recount external events in a way which suggests little emotional or personal significance; the way they think and the way they relate to other people is very predominantly pragmatic. They generally disavow feeling emotions, and so they are also known as “alexithymics” (from the Greek for ‘having no words for emotion’).⁶ However, on the basis of how they interact with other people and the emotions they arouse in others, psychoanalysts argue that they do in fact have emotions, but that they are very out of touch with them. Their form of psychological defense is to “project” their emotions into others. Put simply, they unconsciously imagine that other people have the emotions that they themselves have (while imagining that they themselves do not have these emotions), and interact with them in such a way as to arouse such emotions in others—thus confirming the piece of imagination (see Gardner, 1993, ch. 6; Segal, 1986, ch. 4). In such cases, even at the very time at which such patients undergo an episode of emotional experience, they report having no emotional feelings. It *could be* that such feelings occur in unreflective consciousness and they are simply completely unable to access such feelings. But evidence for this claim is lacking. And we may also argue that in such cases it would be more accurate to say the feelings are unconscious or to argue that no feelings occur at all.

7. The Argument So Far

Let us take stock. We first noted that there is no inconsistency in thinking that unconscious emotions manifest themselves in conscious feelings if we assume that feelings do not reveal the emotions they manifest. However, it is not plausible to think that all cases of unconscious emotion involve simple misunderstanding of conscious feelings, as we are then subject to some very radical misunderstandings indeed. To avoid this objection and explain the lack of understanding, CF1 claimed, contentiously, that the conscious feelings have very little intentional content. To avoid the dilemma, we may adopt Goldie's suggestion that feelings may occur unreflectively: the feeling is a state of consciousness of its object, but the subject is not conscious of the feeling. That such mental states can and do occur is seen in cases of perception in which attention is elsewhere.

CF2 presents a sophisticated interpretation of 'conscious feeling'. Furthermore, it is consistent with the view that psychological defense is driven by mental pain, for it is the *felt* painful awareness of its object that motivates the subject to prevent the feeling from entering reflective awareness and to mask the feeling by others. However, there remain cases, particularly in the psychoanalytic literature, in which the analogy with perception does not apply, and the subject is completely unable, even with effort, to bring her feelings into reflective consciousness. In such cases, it seems reasonable to ask why we should accept that the subject has feelings that are conscious in any sense.

Two paths lie ahead. The first, preserving the claim that feelings are necessarily conscious, argues that emotions—or unconscious emotions at least—need not manifest themselves in feelings at all. The second defends the concept of unconscious feelings.

8. Unconscious Emotions Do Not Involve Feelings 1

8.1. Theory

A number of accounts of emotion in the "cognitive" school defend the view that emotions do not essentially involve "feelings." However, few philosophers have denied outright that there is *any* relation between emotion and feeling. As observed in §2.1, the epistemological access that enables direct and noninferential avowal of emotions is usually thought to involve a type of phenomenology distinctive of emotions. However, cognitivist theories, such as Nussbaum's (2001), usually object to the idea that this is to be understood in terms of "feelings."

Nussbaum argues that emotions are essentially thoughts with a particular type of content (relating to the subject's well-being). She argues, additionally, that there are no extra "noncognitive" elements to emotion, such as feeling, as a correct analysis and understanding of the thoughts involved is sufficient to account for emotion. The distinctive phenomenology, such as it is, can be accounted for by the fact that we

cannot think (episodically) those very thoughts—that the object of the emotion matters to me in a particular way—and retain our equanimity: “The recognizing and the upheaval, we want to say, belong to one and the same part of me, the part with which I make sense of the world” (p. 45).

She then observes that this creates a problem when accounting for (episodic) unconscious emotions, since “if we are prepared to recognize nonconscious emotional states . . . then we cannot possibly hold to any necessary phenomenological condition for that emotion-type” (p. 61). There can’t be any unconscious feelings, since then “we seem to have lost our grip on the notion [of feeling] itself. Is it a kind of psychic energy? But what kind?” (p. 62, footnote). However, this means that to allow for unconscious emotions at all, as she wishes to do, Nussbaum must weaken her claim: “The upheaval is a part of the experience of what it is like to have those thoughts—at least much of the time” (p. 62).

8.2. Evaluation

But it is difficult to see how she can, in consistency, add this qualification. Is the “upheaval” *constituted by* having certain thoughts? Or is it a normal *result*? The most charitable reading is that, when it occurs, the upheaval is constituted by the occurrence of the thoughts; but that *under special circumstances*, the thoughts can occur without the upheaval. In this way, it is the absence of upheaval that requires something additional, not its presence. In these circumstances, the thoughts, and so the episode of emotion, can be unconscious. But we are left in need of some account that explains the separation of the thoughts from the upheaval.

To return to Goldie’s original analysis, we allowed that emotions are dispositions to a variety of episodic mental states and processes, among which are episodes of emotional feeling. This is consistent with claiming that, under certain circumstances, the particular disposition to *feeling* (or upheaval) is never actualized. Let us say that repression, or other forms of psychological defense, undermine or overpower this dispositional property of emotions when rendering them unconscious. The emotions still retain their other dispositional powers, interacting with other mental states in particular ways, motivating us to act, and so on. This solves the puzzle, and we thereby also formulate an informative definition of unconscious emotion—viz., an emotion whose dispositional force to manifest itself in consciousness (i.e., to be felt or cause upheaval) is prevented from achieving fulfillment. A version of this suggestion lay at the heart of Freud’s (1915/1984, §3) theory of unconscious emotion, maintaining that the main aim and achievement of repression is to suppress the development of feeling.

However, we need to consider more closely the episodes of unconscious emotional activity. What is it for an episode of hatred to occur but not be felt? Nussbaum may claim that it is for the relevant evaluative thoughts to occur unconsciously, without upheaval. This may sometimes be true, but the Rat Man *does* appear to be undergoing some kind of “upheaval,” one that Freud traces to the activity of his unconscious hatred. If upheaval is equivalent to “feeling,” for such cases in which the

unconscious emotional activity does involve upheaval, Nussbaum's theory collapses into either the claim that the Rat Man feels his hatred consciously (on the grounds that he undergoes upheaval) or that he feels his hatred unconsciously (on the grounds that the upheaval undergone is not consciously experienced in relation to thoughts of hatred). This is no longer the claim that unconscious emotions occur without feeling.

9. Unconscious Emotions Do Not Involve Feelings 2

9.1. Theory

Roberts (2003) presents a different account of how we may separate emotion, by which he means (in our terms) an episode of emotional activity, from feeling. He accepts an evaluative account of emotional content (emotions are "concern-based construals"; ch. 2), but provides a different account of feelings. To feel an emotion is an "immediate and quasi-perceptual grasp of oneself as in a certain emotional state" (p. 318). Now, "because emotions are self-involving in being based on some concern of the subject, consciousness of the object of the emotion powerfully *predisposes* the subject to be conscious of himself as in the emotional state" (p. 320). Furthermore, like Nussbaum, Roberts emphasizes this must be taken as the norm in the analysis of emotion: "when one *does* feel an emotion, the feeling and the emotion are two aspects of one mental state, rather than two separate ones" (p. 322).

However, this consciousness of oneself is not necessary, and in cases of psychological defense, it doesn't occur: "Emotions are paradigmatically felt, but emotions may occur independently of the corresponding feeling" (p. 60). As noted in §6, we can add that psychological defense may bring about feelings that mask the emotion defended against; e.g., unconscious fear of failure may cause conscious feelings of superiority and even conscious fears of other kinds, but not of failure. Feeling one's emotions, being a "quasi-perceptual" state, like other forms of perception, comes in degrees not only of accuracy, but also of awareness. Psychological defense may work on both accuracy and awareness.

9.2. Evaluation

This line of thought provides a synoptic account of the many forms of "being unconscious" an emotion may take. It allows not only that we may misunderstand what we feel (§5.1), and that we may not be fully aware of what we feel (§6.1), but also that episodes of emotional activity (construals) can occur outside awareness altogether. In such a case, they occur without any feeling at all. The Rat Man simply does not have any "quasi-perceptual grasp" of himself in a state of hatred. Nevertheless, he is, and this fact has consequences for his actions and his other mental states.

This model is suggested by a different motif in psychoanalytic phenomena, viz., that psychoanalytic therapy can acquaint subjects with their emotions for the first time.

As they come to understand the emotions they had during earlier episodes of their lives, they come to feel the emotion now—but this is *the first time* they actually feel the emotion, *pace* Goldie; they have no recollection of feeling during those earlier episodes. This accounts well for the case of alexithymic patients: it is right to say that they simply do not feel their emotions. If psychoanalysis improves their condition, they begin to feel their emotions.

It is not clear, however, that Roberts' (2003) account of feeling can completely resolve our puzzle. Roberts means to connect the verb 'to feel', of which he gives the epistemological analysis above (consciousness of oneself in a particular emotional state), and the phenomenological noun 'feeling'. But I do not think he intends this to be reductive; e.g., he describes the "affect" of an emotion as its "mood," and says that "an important part of the 'feel' of an emotion is its mood" (p. 114). And it is on this phenomenological aspect of feeling that we may press Roberts: when one undergoes an emotion that one does not feel in Roberts' terms (as a construal of oneself), is there nevertheless "something it is like," some "feeling," attached to the episode? Roberts' solution only works if we accept that emotional thoughts can occur without phenomenology. Of any theory which maintains that feeling is distinct from episodes of emotional activity, we may press the question of how the conscious upheavals the Rat Man undergoes are to be explained by a *phenomenologically quiescent* unconscious emotional state. This is a version of the objection made in §5.2.2—whether psychological defense can be properly understood without invoking unconscious feelings.

In what follows, both the psychoanalytically inspired defense of unconscious feelings and the physiological and neurological accounts of feeling seek to provide an analysis of feeling that is not dependent on consciousness, on the basis of more general theories of mental functioning. There is insufficient space to evaluate the theories on which they rely, so I shall not do more than raise general issues regarding the assumed theoretical approach.

10. Unconscious Emotions Involve Unconscious Feelings

10.1. Theory

Gardner (1993) presses the objection repeated above, that we have good reason to believe in *unconscious* feelings, because only if unconscious mental states are *felt* can we explain what we do with them. Only the painfulness of thoughts or emotions explains why and how we reject them. The alternative theory, that the thought or emotion *would become* painful if it were conscious, supposes that consciousness in some sense *creates* the pain (which cannot exist without consciousness), rather than consciousness being *of* something that is, already, painful (p. 216).

To support the argument, we must turn to a general theory of the mind. In talking of emotional feeling, we are talking of the phenomenology of an emotional episode, what it is like, its experiential quality. Wollheim (1984, ch. 2) argues, and many philosophers of mind now agree, that the experiential quality of a mental state and its

representational content are inextricably intertwined (e.g., Crane, 1998; Dretske, 1995; Harman, 1990; Tye, 1995).⁷ In this, Nussbaum is right. According to Wollheim, it is in virtue of the product of the two together that a mental state has the causal powers it does, and so engages in the mental processes it does. It could not have the causal powers it does without the phenomenology it has. Its phenomenal properties, however, are not dependent for their existence on being apprehended in consciousness, and may exist outside consciousness.

This commitment to unconscious phenomenology, unconscious feelings, is not a commitment, Gardner argues, to *unfelt* feelings. Gardner's (and Wollheim's) argument at this point is dense, but my understanding is that feeling is interpreted not as a form of conscious awareness, but in terms of a type of "impact" feelings have (perhaps in contrast with quiescent types of thought). The feelings still impinge upon the subject. The episode of emotional activity has phenomenology, and it has its effects as a result of this phenomenology, i.e., how it feels. But the subject may not be conscious of this feeling: "What it is that extends beyond consciousness in the case of unconscious pain is just what it is that is apprehended when pain is consciously given. Thus it would be a mistake to think that the supposition of unconscious pain involves *unfelt* or *unfeeling* pain" (Gardner, 1993, p. 217). We may extend the point to emotional feelings generally. Gardner's analysis of the Rat Man is, then, that he feels his hatred unconsciously, and this episode of feeling interacts with other mental states and events, such as his motivation to replace the stone, in ways similar to how it would were it conscious (allowing for the different nature and modes of expression of unconscious states noted by psychoanalytic theory).

Although Gardner's analysis sounds similar to Goldie's argument for unreflective feeling, it is worth noting the difference between the two accounts. The model of unreflective awareness discussed in §6, and the arguments by which it was set up, suggested that emotional feelings are *available* to the subject's reflective consciousness, by means of redirecting attention, even if this requires effort to overcome the motivation not to recognize the existence or nature of the feelings. In Gardner's hands, unconscious feeling can be completely inaccessible to consciousness. There is no suggestion here that we may be able to access to it by directing our attention in a certain way. Goldie ties feeling to consciousness far more closely than Gardner, arguing for a separation only as a result of inattention. We therefore have a three-fold division: what is reflectively conscious, what is unreflectively conscious, and what is unconscious. However, it is worth noting that what is unconscious in Gardner's sense is still dependent on consciousness in general. No organism that cannot feel consciously can feel unconsciously—the psychoanalytic unconscious occurs only in beings with relatively sophisticated consciousness.

10.2. Evaluation

Unsurprisingly, this theory is the most potent of those we have surveyed in its explanatory force: there are no cases that make it fail as a general account of unconscious emotion. It does not need to insist that *every* case of what we may rightly

term unconscious emotion must be interpreted in terms of unconscious feeling, for there is no threat to it in allowing that the explanations offered in §§5.1 and 6.1 may hold true in certain cases. Its philosophical weakness, many will say, lies in its commitment to unconscious phenomenology.

Unconscious feelings are not to be understood by analogy with cases of hearing a drill in the background or coming to after a period of absent-minded driving. But for precisely this reason, independent of the theoretical considerations, it can be difficult to accept that unconscious states, in contrast to reflectively and unreflectively conscious states, have phenomenology. That there can be no unconscious feelings is still the position of “commonsense.” But this is not a conclusive objection: commonsense is what produces the puzzle of unconscious emotion in the first place, and Gardner and Wollheim’s general theory of mental functioning, which supports Gardner’s theory of unconscious emotion, is, they argue, an extension of commonsense psychology (Gardner, 1993, ch. 1, 4; Wollheim, 1993, ch. 6; see also Hopkins, 1982). If we can find sufficient explanatory drive to posit them, then we have reason to accept their existence. The need to appeal, in cases like the Rat Man and alexithymic patients, to phenomenological states to which subjects cannot turn their attention provides the requisite explanatory demand.

Given the general difficulty here of accepting the idea of unconscious phenomenology, despite the disanalogy we have remarked between unconscious feelings and perceptual cases, defenders of unconscious feeling may take strength from the fact that a number of theorists have independently defended the claim that phenomenal properties can occur independently of the subject’s being conscious of them (e.g., Burge, 1997; Rosenthal, 1991). For example, Burge’s defense is not dissimilar to Gardner’s, arguing that while what it is like to experience a phenomenal property in consciousness is essential to typing that property, the property may occur without consciousness (pp. 432ff.). He also allows that not all such cases of unconscious phenomenology are simply a matter of inattention. Of course, this is not uncontroversial, but if we need to accept nonconscious phenomenology quite independently of psychoanalytic considerations, this supports Gardner’s hand. However, critics will reply that the implications of the debate in perception are only that *unreflectively conscious* states have phenomenology, so that it is Goldie’s theory, not Gardner’s that gains support.

11. Feelings and Physiology

11.1. Theory

As a final attempt to remove some of the “mystery” of what unconscious feelings (or phenomenal properties generally) could be, it can be tempting to provide an account of them in nonpsychological terms. If feelings could be identified with some physiological or neurological state, then we could have criteria independent of the subject’s consciousness for asserting that such feelings do or do not exist. The first path, that of identifying feelings with physiology, has a long history in the theory of

emotion; the second is a more recent development. Damasio (1994) provides a way into the first, his (2000) into the second.

Damasio (1994, especially ch. 7) argues that the essence of an emotion, by which he means an episode of emotion, is changes in one's *body*, such as pulse, breathing, skin conductance but also visceral and hormonal alterations, in response to an evaluative image of the emotion's object. The profile of these changes is then represented in the brain. However, Damasio then argues that a subject does not *feel* an emotion unless she experiences the bodily changes and relates them to the images which caused them. An episode of emotional activity and feeling an emotion are distinct.

Each family of theories of unconscious feelings could attempt to use Damasio's ideas to understand unconscious emotions. The *conscious feelings* family could argue that a subject will be (unreflectively) aware of, but misunderstand, the physiological changes that occur. Without relating the experienced bodily changes back to the object of the emotion, to the extent that she is aware of them, the subject doesn't understand the feelings as emotional (they are experienced as bodily feelings instead, e.g., heart palpitations, not fear). The *no feelings* family could emphasize Damasio's claim that feeling an emotion is distinct from undergoing an emotional episode, arguing that bodily changes provide us with a criterion for an episode of emotional activity without feeling. The *unconscious feelings* family could adapt the theory to claim that the bodily changes comprising the emotional episode provide us with an account of the ontology of unconscious feelings: these bodily changes have the requisite sort of "impact" on the subject to count as feelings, all that is missing is consciousness of them as such (a significant emphasis on the involvement and role of the body in unconscious emotion is uncontroversial in most psychoanalytic theories).

11.2. Evaluation

Although many theories are happy to accept that feelings are, in some sense, grounded in or related to (if not reducible to) physiological alterations, there are two important objections to using the theory to resolve the puzzle of unconscious emotions. First, it is unclear whether such bodily changes could tell us *which* emotion the subject is undergoing. Given that the Rat Man experiences a number of conscious feelings, could his physiological state alone also determine whether he was also undergoing hatred? Cases of episodes of unconscious emotional activity often involve multiple emotions, and there is currently little evidence to suggest that evidence from physiological responses could make the requisite distinctions even in far less complicated cases (see Prinz, 2004, pp. 72–74, for discussion). Second, Damasio goes on to allow that changes in the body needn't actually occur for feeling to take place. It is possible for the mental images that would trigger the physiological changes in the body to instead trigger an "as if" representation of the body state in the brain, and this can be sufficient for feeling. It is possible, therefore, for a subject to undergo a felt episode of emotion without changes to the body. If such an "as if" representation of

body state could occur without consciousness, there could there be unconscious episodes of emotion—indeed, unconscious feelings—that do not involve bodily changes. So physiology will not provide a reliable criterion for the occurrence of unconscious emotional episodes.

12. Feelings, Physiology, and Phenomenal Properties

12.1. Theory

When, in his later work, Damasio (2000) discusses the question of feelings and consciousness (pp. 279–285), he distinguishes between “having a feeling”—essentially the representation of the bodily changes that constitute the emotion (whether those changes in fact occur or are represented “as if”)—and “feeling a feeling,” a matter of knowing that one has that feeling, which involves relating the feeling to the sense of self. This entails that feelings can occur without one’s knowledge of them, i.e., without one’s consciousness of them. In this development of the theory, however, the possible existence of unconscious feelings is not tied to physiological changes, but to mental representations that “arise out of” patterns of neural firing.

12.2. Evaluation

I am unclear whether Damasio wishes to endorse unconscious feelings. He endorses very extensive unconscious mental processes (pp. 226–228, 296–302), but in the discussion above, he states: “When those images [that comprise feeling an emotion] are accompanied, *one instant later*, by a sense of self in the act of knowing... they become conscious” (p. 282, italics added). Whether, in organisms that have the “core consciousness” (ch. 3, 6) that grounds the requisite sense of self, feelings could ever *remain* unconscious, even if they begin that way, is unclear. It seems that Damasio could equally endorse a position like Goldie’s; or again, given the importance of relating the emotion to the self in the production of feeling, there are similarities to Roberts’ “no feeling” theory. Damasio’s neurophysiological theory does not settle which solution to our problem we should settle on.

A number of philosophers have sought to identify phenomenology with neurophysiological properties (e.g., Kirk, 1994; Prinz, 2004; Tye, 1995). To defend unconscious feeling against the commonsense objection to unconscious phenomenology of any kind, one could bring Damasio’s model together with these more general philosophical accounts: an unconscious feeling does occur, because the requisite neural pattern occurs.

This move faces two problems. First, as remarked at the end of §10.2, the philosophical accounts are developed for perceptual phenomenal properties that are *readily accessible* to consciousness, although not in fact conscious. Second, the identification of phenomenal properties with properties of brain functioning is hugely controversial. Damasio (2000) wants to hold a distance between them, saying

the mental properties “arise out of” the neural ones (Appendix), but that the latter could, and do, occur without the former (e.g., p. 74). Unless we take issue with this claim, we still have little reason for thinking that the occurrence of the neural property in the absence of consciousness is tantamount to an occurrence of unconscious phenomenology rather than the absence of phenomenology. So although this line of argument may seem to offer support to Gardner’s defense of unconscious feelings by providing a realist ontology, Gardner (1993) himself claims it is at odds with his explicit attempt to preserve the *painfulness* of unconscious pain, not merely its neurological underpinnings (p. 218).

13. Conclusion

This survey has not reached any firm conclusions. Within the conscious feelings family, philosophers may insist that feelings must, in some sense, be conscious. To the objection that without subjects’ recall, there is no evidence for their position, they may now use the argument of the unconscious feelings theorist: episodes of unconscious emotional activity are not phenomenologically quiescent, but have an “impact” on the subject. The no feelings family may reply that no feelings occur, and the proper of understanding this “impact” is not through phenomenology but through the thought content of the emotion (this response is open to Roberts, but not Nussbaum, as she equates thinking the thought content to upheaval). As remarked at the outset, emotions are manifest in various effects upon other mental states; unless one accepts Wollheim’s theory of mental functioning, according to which phenomenology is necessary for such effects, the no feelings theorist may argue that they occur in the absence of feelings. Finally, the unconscious feelings theorist may argue that this commitment to equating feelings to conscious feelings is undermined by explanatory requirements emerging both from discussions of perception and from psychoanalytic case studies. Following Freud’s (1915/1984) distinction between the preconscious and the unconscious, they may further argue that there is a distinction *in kind* between that which prevents unreflectively conscious feelings from entering (reflective) consciousness and that which prevents unconscious feelings from so doing, and so the theory of unreflective consciousness cannot do the requisite work. All three theories may (but need not) appeal to Damasio’s work on the bodily origins and/or nature of emotional feelings to help in explaining the “impact” that unconscious emotions may have.

Some philosophers, drawing on considerations from Wittgenstein or Merleau-Ponty, may feel that the discussion took a wrong turn—“inwards”—in §10. Rather than defend unconscious feelings via a realist theory of “phenomenal properties” (whether or not these are identified with neural properties), the question should be even more firmly situated within the need to explain human behavior. The development of a metaphysical mental realism will not solve the question of what the Rat Man felt nor what alexithymics feel nor explain what role the notion of “unconscious feeling” can play in our understanding of human life.

Many psychoanalysts will note that more needs to be said about the ways in which psychological defense can *transform* emotions and emotional feelings (e.g., Lacewing, 2005), as in the case of alexithymics' projection of emotion. I confess some sympathy with both these responses, and suspect that future work in developing a solution to our puzzle will need to take account of them.

Notes

- [1] This debate arises in §5, where I give some support to the claim that feelings are not just qualia, i.e., entirely divorced from the intentional content of the emotion.
- [2] Psychoanalysts might well add to this an account of the history of an emotion in terms of fantasy. Fantasy may be seen as the vehicle by which our present emotional experience is coloured by our past emotional objects; the history of emotion is the history of fantasy and its objects. For a philosophical exploration and defence, see Gardner (1992, 1993, ch. 6).
- [3] This claim needs to be distinguished from the different and separate claim that the feelings may be indistinct because the emotion itself is *indeterminate*. Charles Taylor (1985) argues that our emotions involve self-interpretation, and it may not be true of me that I have a particular emotion until I have interpreted myself as having that emotion. This is not a claim that is particularly about unconscious emotion, for once I have reached my self-interpretation, supposing it is sufficiently accurate and true to myself to resolve my uncertainty, the resulting determinate emotion is conscious. I am supposing that in cases such as the Rat Man, that he feels hatred—even if the precise content of that hatred is indeterminate—is sufficiently determinate independent of and prior to his self-interpretation. Freud's explanation—and indeed, I believe, any plausible explanation—of the stonemoving episode requires this degree of mental realism. That Taylor (1985) would probably accept this realism as consonant with his interpretive constructivism is indicated by his acceptance of limits on self-interpretation: “in offering a characterization [of our feelings], these feelings open the question whether this characterization is adequate . . . whether we have properly explicated what the feeling gives us a sense of” (p. 64). (See also Tanney (2002)).
- [4] I thank an anonymous reviewer for clarifying the need to present and defend my objections in detail, and for a number of the points that follow, defending the theory that unconscious emotions involve conscious feelings.
- [5] The *locus classicus* on defense mechanisms is Anna Freud (1968). For a brief introduction, see Lacewing (2005). For a lengthier philosophical discussion, see Gardner (1993, ch. 5–6).
- [6] If an analogy with perception is to be made, it might seem that blindsight would provide a better counterpart, and Prinz (2004, ch. 9) explores this possibility in relation to alexithymia. However, despite the analogy that the subject denies conscious perceptual/emotional experience, there are many important disanalogies regarding forced guessing and behaviour guidance.
- [7] Rather confusingly, Wollheim calls the experiential quality ‘subjectivity’ rather than ‘phenomenology’, and uses ‘phenomenology’ to refer to the joint product of subjectivity and intentionality. I follow the use of ‘phenomenology’ that is more common in the philosophy of mind literature.

References

- Armstrong, D. (1968). *A materialist theory of mind*. London: Routledge & Kegan Paul.
 Ben-Ze'ev, A. (2000). *The subtlety of emotion*. Cambridge, MA: MIT Press.

- Block, N. (1997). On a confusion about a function of consciousness. In N. Block, O. Flanagan, & G. Güzelde (Eds.), *The nature of consciousness: Philosophical debates* (pp. 375–415). Cambridge, MA: MIT Press.
- Burge, T. (1997). Two kinds of consciousness. In N. Block, O. Flanagan, & G. Güzelde (Eds.), *The nature of consciousness: Philosophical debates* (pp. 426–433). Cambridge, MA: MIT Press.
- Clore, G. (1994). Why emotions are never unconscious. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 285–290). New York: Oxford University Press.
- Crane, T. (1998). Intentionality as the mark of the mental. In A. O’Hear (Ed.), *Current issues in philosophy of mind* (pp. 229–251). Cambridge, England: Cambridge University Press.
- Damasio, A. (1994). *Descartes’ error*. New York: G. P. Putnam’s Sons.
- Damasio, A. (2000). *The feeling of what happens*. London: Heinemann.
- Descartes, R. (1989). *The passions of the soul* (Stephen H. Voss, Trans.). Indianapolis, IN: Hackett. (Original work published 1650)
- Dretske, F. (1993). Conscious experience. *Mind*, 102, 263–283.
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Freud, A. (1968). *The ego and the mechanisms of defence* (Rev. ed.). London: Hogarth Press.
- Freud, S. (1979). *The Penguin Freud library: Vol. 9. Notes on a case of obsessional neurosis*. London: Penguin. (Original work published 1909)
- Freud, S. (1984). *The Penguin Freud library: Vol. 11. The unconscious*. London: Penguin. (Original work published 1915)
- Gardner, S. (1992). The nature and source of emotion. In J. Hopkins & A. Savile (Eds.), *Psychoanalysis, mind and art* (pp. 35–54). Oxford, England: Blackwell.
- Gardner, S. (1993). *Irrationality and the philosophy of psychoanalysis*. Cambridge, England: Cambridge University Press.
- Goldie, P. (2000). *The emotions*. Oxford, England: Oxford University Press.
- Greenspan, P. (1988). *Emotions and reasons*. New York: Routledge.
- Harman, G. (1990). The intrinsic quality of experience. In J. Tomberlin (Ed.), *Action theory and philosophy of mind* (pp. 31–52). Atascadero, CA: Ridgeview.
- Hopkins, J. (1982). Introduction: Philosophy and psychoanalysis. In R. Wollheim & J. Hopkins (Eds.), *Philosophical essays on Freud* (pp. vii–xlv). Cambridge, England: Cambridge University Press.
- Kirk, R. (1994). *Raw feeling*. New York: Clarendon Press.
- Lacewing, M. (2005). Emotional self-awareness and ethical deliberation. *Ratio*, 18, 65–81.
- McDougall, J. (1986). *Theatres of the mind*. London: Free Association Books.
- Nussbaum, M. (2001). *Upheavals of thought*. Cambridge, England: Cambridge University Press.
- Prinz, J. (2004). *Gut reactions*. New York: Oxford University Press.
- Roberts, R. (2003). *Emotions: An essay in moral psychology*. Cambridge, England: Cambridge University Press.
- Rorty, A. (1980). Explaining emotions. In A. Rorty (Ed.), *Explaining emotions* (pp. 103–126). Berkeley, CA: University of California Press.
- Rosenthal, D. (1991). The independence of consciousness and sensory quality. In E. Villanueva (Ed.), *Consciousness: Philosophical issues* (Vol. 1, pp. 15–36). Atascadero, CA: Ridgeview.
- Segal, J. (1986). *Phantasy in everyday life*. London: Penguin Books.
- Tanney, J. (2002). Self-knowledge, normativity, and construction. In A. O’Hear (Ed.), *Language* (pp. 37–55). New York: Cambridge University Press.
- Taylor, C. (1985). Self-interpreting animals. *Human agency and language* (pp. 45–76). Cambridge, England: Cambridge University Press.
- Tye, M. (1995). *Ten problems of consciousness*. Cambridge, MA: MIT Press.

Wollheim, R. (1984). *The thread of life*. Cambridge, England: Cambridge University Press.

Wollheim, R. (1993). *The mind and its depths*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1999). *On the emotions*. New Haven, CT: Yale University Press.

Wollheim, R. (2003). The emotions and their philosophy of mind. In A. Hatzimoysis (Ed.), *Philosophy and the emotions* (pp. 19–38). Cambridge, England: Cambridge University Press.