

# Sentimentalism and Moral Dilemmas

András SZIGETI

UiT The Arctic University of Norway/Linköping University

Email: [andras.szigeti@uit.no](mailto:andras.szigeti@uit.no)

**ABSTRACT:** It is sometimes said that certain hard moral choices constitute tragic moral dilemmas in which no available course of action is justifiable, and so the agent is blameworthy whatever she chooses. This paper criticizes a certain approach to the debate about moral dilemmas and considers the metaethical implications of the criticisms. The approach in question has been taken by many advocates as well as opponents of moral dilemmas who believe that analyzing the emotional response of the agent is the key to the debate about moral dilemmas. The metaethical position this approach is most naturally associated with is sentimentalism. Sentimentalists claim that evaluation, and in particular moral evaluation, crucially depends on human sentiment. This paper is not concerned with the question whether moral dilemmas exist, but rather with emotion-based arguments used on both sides of the debate. The first aim of the paper is to show that emotion-based arguments by friends or foes of moral dilemmas cannot garner support from sentimentalism. The second aim is to show that this constitutes a serious problem for sentimentalism.

## 1. Introduction

This paper criticizes a specific kind of argument in the debate about moral dilemmas and considers the metaethical implications of this critique. Moral dilemmas are special moral conflicts in which no available course of action is justifiable and therefore the agent is blameworthy<sup>1</sup> whatever she chooses. Much of the philosophical debate has focused on whether such situations are possible.<sup>2</sup>

Many believe that the emotional response of the agent facing such a situation could be the key to resolving this debate. It is frequently observed that agents in some conflict situations will typically feel guilt whatever they choose. It is then added that the agent's guilt appears apposite. However, guilt is usually taken to be apposite if and only if the action is unjustifiable. It would follow that moral dilemmas exist.

This paper is not concerned with the question whether moral dilemmas exist or not. Rather, what it focuses on is the appeal to emotions in arguments about moral dilemmas such as the one above. Emotion-based arguments, as I shall call them, presuppose that in order to properly evaluate dilemmatic situations we need to focus on our characteristic emotional responses to such situations.

The metaethical view this presupposition is most naturally associated with is sentimentalism. Sentimentalists claim that "evaluation, and in particular moral evaluation, is somehow grounded in human sentiment" (D'Arms and Jacobson 2000b, 722). How exactly value is grounded in human sentiment can be spelled out in different

---

<sup>1</sup> I will use the term "blameworthy" throughout in a non-affective sense to refer to wrongdoings for which the agent has no justification and no responsibility-undermining excuse. According to this definition, a blameworthy action is an all-things-considered unjustifiable action for which the agent is morally responsible. Note that on this understanding of blameworthiness: "A person can be praiseworthy or blameworthy without anyone's being aware of this, without anyone's taking note of it, without anyone's actually praising or blaming him" (Zimmerman 1988, 39).

<sup>2</sup> It is worth noting that the term "moral dilemma" is frequently used somewhat misleadingly to designate a much wider range of hard moral choices (see, for example, Greene et al., 2001). In order to focus the issue, it is advisable to reserve the term only to those special moral conflicts (if any) in which all available courses of action are unjustifiable/blameworthy. The term "moral conflict" can then be used for situations in which (i) there is an all-things-considered justified course of action, but (ii) choosing that course of action has a high moral cost, for example, because it entails not discharging some duty or violating somebody's basic rights. Finally, according to this terminology "moral quandaries" could be defined as situations involving special epistemological obstacles to ascertaining the right course of action. Many moral dilemmas and moral conflicts are *not* moral quandaries in this sense (Williams, 1973 is especially clear on this point).

ways, but sentimentalism does appear to provide the right sort of theoretical justification for relying on emotion-based arguments. Dilemmatic situations should be understood by way of emotional responses to them because in general “evaluation is to be understood by way of human emotional response” (D’Arms and Jacobson 2005, 187-188).

Conversely, a plausible emotion-based dilemma-argument, whether for or against, could increase the attractiveness of sentimentalism. The success of such an argument could help to show that we must indeed call on emotions as the final court of appeal for the appropriateness of our evaluations. We would obtain additional support for the view that the appropriateness of certain emotions constitutes a decisive criterion for the ascription of evaluative properties.

The first aim of the paper is to show that emotion-based arguments by friends or foes of moral dilemmas cannot garner support from sentimentalism. The second aim is to show that this is not a marginal issue but constitutes a serious problem for sentimentalism. After the presentation of what is by now a classic example (Section 2), I will argue for these negative conclusions in the middle part of the paper (Sections 3-7). Section 8 concludes.

## 2. Sophie’s Choice

*Sophie’s Choice* (Greenspan, 1983), many argue, is the purest example of a genuine moral dilemma. Sophie receives this offer: “choose between your two children, otherwise both will be killed”. Sophie chooses one of her children and so only one child is killed. “Sophie” of course is just a placeholder name here for a mother with “ordinary moral sensibility”.<sup>3</sup>

The point of departure for friends of moral dilemmas is the observation that Sophie will feel guilty about her choice whatever she chooses: sacrificing one child or letting both die if she rejects the offer. The second observation is that guilt is an appropriate reaction whatever she chooses. Imagine a mother reporting no guilt feelings about having allowed her child (or children) to die. Would we not be deeply troubled by such a reaction? Would we not expect her to feel guilt? What’s more, would we not find the lack of guilt *morally* repugnant (see, for example, Greenspan 1983, 120)? But if that is the case, then it seems that Sophie’s guilt is appropriate. And since lack of guilt would be morally repugnant whatever she were to choose, Sophie’s guilt can be said to be appropriate whatever Sophie chooses.

Further, friends of moral dilemmas can point out that, according to the common understanding of guilt, agents of “ordinary moral sensibility” feel guilt about actions that they cannot justify and for which they are morally responsible. In short, guilt is appropriate if (and only if) the agent is blameworthy. But we have already established that guilt would be appropriate whatever Sophie chooses. So the surprising conclusion is that whatever Sophie chooses to do her action will be unjustifiable and blameworthy. Therefore, she is in a moral dilemma (see Williams, 1973; Marcus 1980; McConnell 1996).

---

<sup>3</sup> The expression is used by Marcus (1980, 198).

I will use this analysis of *Sophie's Choice* as a model argument. It well represents the way emotion-based arguments have been generally used in the moral dilemmas debate.<sup>4</sup> What I am interested in is the appeal to Sophie's *guilt* in this argument. Can it carry the argumentative burden friends of moral dilemmas place on it? I will not question the claim that people with "ordinary moral sensibility," were they in Sophie's shoes, would experience an emotion resembling normal guilt. Rather, the worry is whether existing similarities between guilt proper and Sophie's guilt can license the evaluative conclusion of the emotion-based argument.

Note that Sophie's guilt being "an emotion in the neighborhood" (D'Arms and Jacobson 1994, 743) of guilt is not enough for the emotion-based pro-dilemma argument to go through. Foes of moral dilemmas have also used emotion-based arguments. But they describe Sophie's feelings differently. *Their* point of departure is the observation that what Sophie feels is regret (McConnell, 1978). No doubt, they add, we would be troubled if Sophie did not regret what happened. In fact, we would probably be morally repulsed by the lack of regret. Moreover, such regret would be appropriate whatever Sophie chose to do.

But, and this is the crucial move, regret is different from guilt in that we can appropriately regret doing something that we were justified in doing. It is appropriate to regret that I had to break into my neighbor's flat even if that was the only way I could obtain life-saving medication for a guest of mine.<sup>5</sup> By contrast, feeling guilt would not be apposite precisely because the course of action taken was clearly justified. So although her regret is appropriate whatever she does, it does not follow that whatever Sophie does her action will be unjustifiable and blameworthy. Therefore, she is not in a moral dilemma.

It emerges that quite a lot could depend on whether Sophie's emotional response is best described as guilt proper or rather as an emotion such as regret that only superficially resembles guilt. So how are we to describe Sophie's dominant emotional response?

### 3. What Does Sophie Feel?

It cannot be excluded *a priori* that we will find important behavioral and phenomenological differences between Sophie's emotional response and guilt proper. This is an empirical matter (remember, Sophie could be any agent with "ordinary moral sensibility"). With sufficient observational data in hand, we might indeed conclude that people's behavior and experience in dilemmatic situations differ from normal behavioral and phenomenological patterns associated with guilt.<sup>6</sup> For example, some have speculated that agents in dilemmatic situations characteristically "experience moral doubt" and frequently "seek moral advice" (McConnell 1978, 163-4). There is some reason to be skeptical whether this particular attempt really succeeds in

---

<sup>4</sup> Cases similar to *Sophie's Choice* have been used to construct emotion-based arguments by friends and foes of moral dilemmas alike. Such arguments have been developed, among others, to analyze Sartre's story of a young Frenchman who has to choose between joining the resistance or looking after her mother (see footnote 27), Agamemnon's agony whether to sacrifice her daughter or betray his duty as a commander, or Captain Vere's condemnation of Billy Budd (apart from the *locus classicus* Williams, 1973, see esp. Marcus, 1996; McConnell, 1996; Gowans, 1996; and Mothersill, 1996 – several other contributions to Mason 1996 discuss emotions in moral dilemmas as well). In the specific case of *Sophie's Choice*, it might be argued to raise special complications that Sophie is the children's mother. I will come back to this point.

<sup>5</sup> See also Feinberg's original cabin example (1978, 102) that this scenario is modeled on.

<sup>6</sup> For experimental data on responses to moral conflicts and moral dilemmas, see among others Greene et al. (2001), Koenigs et al. (2007), and Bartels and Pizarro (2011).

identifying unique phenomenological or behavioral features of agential reactions to dilemmatic situations. Thus it is not at all clear that doubt is typical in dilemmatic situations and atypical in non-dilemmatic ones (Williams, 1973).

But let us assume for the sake of the argument that we do manage to find some significant behavioral or phenomenological differences between ordinary guilt and Sophie's emotional response. The crucial question is this: What could we conclude from the observation that such differences exist? Answer: not much.

Let me explain. Any theory of emotion that regards emotions as representational mental states – and the overwhelming majority of emotion theories do – must allow for the possibility of incorrect token emotional reactions. Any mental state must be capable of being incorrect if it is to function as a representation. If emotions represent they must also be able to misrepresent (Prinz, 2004).

This follows from a general truth about mental representations. In addition, we must allow for incorrect emotional reactions for a more specific reason. The overwhelming majority of metaethical theories, sentimentalism and many forms of non-sentimentalist alike, assume that there is an illuminative relationship between value and emotion. For example, emotional reactions such as shame and amusement are supposed to reveal significant facts about what is shameful or funny, respectively. If this is true, then it cannot be the case that absolutely anything that happens to cause shame is shameful or whatever amuses is funny.

Now if incorrect emotional representations are a general possibility, then it cannot be excluded that Sophie's emotional response is such an incorrect reaction. Specifically, it is possible that whatever behavioral or phenomenological differences we find between ordinary guilt and Sophie's emotional reaction these differences are due to the malfunctioning of the agent's emotional sensibilities. So it could be argued that guilt would have been the fitting reaction by Sophie, but for some reason her emotional reaction gets derailed, and so ends up being phenomenologically and behaviorally *different* from ordinary guilt.<sup>7</sup>

Friends of moral dilemmas will want to argue that this is indeed the case. They may or may not add a theory to provide a systematic explanation of why Sophie's emotional reaction can be *expected* to differ from ordinary guilt given some characteristics of Sophie's situation.<sup>8</sup> By contrast, foes of moral dilemmas will argue that regret – or perhaps some other related emotion such as sadness – is the fitting reaction. They might regard the presence of behavioral and phenomenological differences between ordinary guilt and Sophie's emotional response as evidence supporting their view. But, and this is the salient point, it is unclear that the fittingness of regret should be a better explanation of why Sophie's reaction differs phenomenologically and behaviorally from ordinary guilt than the explanation from malfunctioning emotional sensibilities. No matter how accurately we describe the behavior of agents and the phenomenology of their emotional experience in dilemmatic situations, a merely descriptive approach cannot settle which party, the friend or the foe of moral dilemmas, is right.

---

<sup>7</sup> The emotional reaction can fail to be fitting due to the presence of “defeaters” (see Tappolet 2000, esp. 216-226) or “obscuring factors” (D'Arms and Jacobson, 2010) that undermine the normal functioning of the agent's emotional sensibility (e.g., stress, mood effects, peer pressure). These factors may be episodic or systematic, idiosyncratic or common to most human beings. Some of them are also cited as typical causes of perceptual errors.

<sup>8</sup> Such a theory would serve a similar purpose as systematic explanations of well-known perceptual illusions such as, for example, that of the Müller-Lyer illusion.

By the same reasoning, if we were to find *no* significant behavioral or phenomenological differences between ordinary guilt and Sophie's reaction that would not settle the debate about moral dilemmas either. Just as before, it may be that Sophie's response ends up being behaviorally and phenomenologically indiscernible from ordinary guilt phenomenology and behavior because of a malfunctioning of Sophie's emotional sensibilities. Foes of moral dilemmas will want to argue that this is indeed the case and that regret – or some similar emotion such as sadness – would be the fitting reaction. Once again, foes of moral dilemmas could also attempt to supply a systematic explanation of why, given some characteristics of Sophie's situation, Sophie's emotional reaction can be expected to be indiscernible from ordinary guilt despite guilt not being the fitting response in this particular case. By contrast, friends of moral dilemmas think guilt would be the fitting reaction, and could add that the lack of behavioral and phenomenological differences supports their view. Sophie's guilt is phenomenologically and behaviorally indiscernible from ordinary guilt precisely because it is the fitting reaction, they could say. But, again, it is unclear that fittingness of guilt should be a better explanation of why Sophie's guilt is phenomenologically and behaviorally indiscernible from ordinary guilt than the explanation from malfunctioning emotional sensibilities.

An additional reason to question the significance of behavioral and phenomenological differences or similarities between ordinary guilt and Sophie's emotional reaction is the so-called "Conflation Problem" insightfully analyzed by D'Arms and Jacobson (2000a; 2000b; 2003). The problem, in short, is that in many cases we can have good reasons not to respond with the emotion that the correct evaluation of the situation would render fitting. For example, it could sometimes be morally or for some other reason inappropriate to be amused by a joke (e.g., at a funeral). Still, that does not mean that amusement is not a fitting response if the joke is funny.<sup>9</sup> It is not unrealistic to assume that we sometimes bring ourselves, deliberately or relenting to unconscious pressures, to feel certain emotions (especially, if we are morally expected to feel those emotions). So Sophie may feel whatever she feels because there are such reasons in play – good, respectable reasons but the wrong kind of reasons nevertheless, since they do not bear on what emotional response her situation would render fitting. For example, quite a few people, including Sophie herself, might think that Sophie ought to feel guilt because, say, she owes at least that much to the memory of her dead child, or because in such cases it is better to err on the side of too much rather than too little guilt.<sup>10</sup> But the availability of such reasons, even moral reasons, does not necessary imply that Sophie's guilt is fitting in her particular situation.

In sum, the mere fact (if it is a fact) that Sophie's guilt and ordinary guilt are indistinguishable in phenomenological or behavioral terms should not be seen as evidence supporting the emotion-based pro-dilemma argument. By the same token, the mere fact (if it is a fact) that there are significant phenomenological or behavioral differences between Sophie's emotional response and ordinary guilt should not be seen as evidence supporting the emotion-based anti-dilemma argument. Behavioral and phenomenological differences or similarities between Sophie's guilt and ordinary guilt

---

<sup>9</sup> This leads to a general difficulty – not just for sentimentalists but all adherents of the fitting-attitudes analysis of value – how to distinguish between reasons for an attitude due to the fittingness of the evaluation the attitude represents, and wrong kind of reasons that have to do with the moral or instrumental appropriateness of the attitude (Rabinowicz and Rønnow-Rasmussen, 2004).

<sup>10</sup> For the latter kind of explanation of why Sophie ought to feel guilt, see Dahl 1996, 95: "For example, it may be that people can't achieve a fine enough control over their feelings and emotions that, if they are disposed to feel guilty in situations in which they have done something wrong and are at fault, they won't be able to prevent themselves from feeling guilty in comparable situations in which they are not at fault."

are inconclusive by themselves because they leave it open whether Sophie's emotional reaction, whatever it may be, is fitting or not.

#### 4. Fitting Emotions

How do we then settle the question about the fittingness of Sophie's emotional response? In order to answer this question, we first need to address the wider problem of what makes guilt a fitting reaction in general.

Since guilt is an emotion and emotions are representations of evaluative concerns, the fittingness of guilt will depend on what guilt is *about*, i.e., the evaluative concern that guilt represents. We feel guilty about a great number of things: for offering and accepting bribes, for traffic violations, marital infidelities, murders, war crimes, our forgetfulness, and tax evasions. But if guilt is to qualify as a cohesive emotion *type*, there must be something common to these episodes. At the very minimum, what they all share is that their object is represented by the subject of the emotion as "guilt-worthy". If so, then the fittingness of someone's guilt depends on whether what this person feels guilty about is "guilt-worthy" or not. That is, guilt is fitting if the object of guilt is such that it warrants or calls for or merits guilt.<sup>11</sup>

In general, the concept of "core relational theme" (Lazarus, 1991) is a useful way to capture what is shared by all token episodes of a given emotion type.<sup>12</sup> Each such theme glosses the distinct way in which that emotion presents the world to its subject.<sup>13</sup> Now, all we have said so far about the core relational theme of guilt is that guilt represents the "guilt-worthy". Of course, for the notion of core relational theme to be informative we need to say more about what is common to all instances of an emotion type than that they represent their object as "worthy" of that type of emotional reaction.

There have been many attempts to provide illuminative characterizations of the core relational themes of various emotions. For example, it is said that the core relational theme of sadness is "irrevocable loss" (Prinz and Nichols 2010, 119). On the whole, approximately, the "core relational theme" circumscribes the evaluative property the emotion tracks, e.g., the property of being "irrevocably lost" for sadness, or the property of being "funny" for amusement, "threatening or frightening" for fear, and so on.<sup>14</sup>

Here are a few more suggestions for the core relational themes of various emotions:

*"Jealousy monitors the social environment for potential losses of affection or allegiance... Its characteristic appraisal is... defection."* (D'Arms and Jacobson 2003, 139)

*"One's envy portrays a rival as having a desirable possession that one lacks, and it casts this circumstance in a specific negative light."* (D'Arms and Jacobson 2000a, 66)

---

<sup>11</sup> For the purposes of this paper, we can leave the metaethical status of the standard of warrantedness open, i.e., we need not specify whether it is intrapersonal or interpersonal. This issue divides different varieties of sentimentalism. See note 21 below.

<sup>12</sup> Talk about the "formal object" of a given emotion type serves the same purpose (Teroni, 2007).

<sup>13</sup> Some believe that representation must take the form of propositional judgment (Lazarus, 1984). Others reject this requirement denying that the representation of the evaluative stance must be propositional or even conscious in the affective episode (Tappolet, 2000; Prinz and Nichols, 2010; Döring, 2010).

<sup>14</sup> Why only approximately? Because it can be argued (see esp. Mulligan, 2009) that several distinct evaluative properties are tracked by the same kind of emotion (admiration tracks generosity, charm as well as elegance), and the same evaluative property can be tracked by different emotions (both admiration and pride track elegance).

“So reasons of fit for *fear* are roughly those that speak to whether or not something is a *threat*.” (D’Arms and Jacobson 2006, 108)

“*Shame* presents something as a social disability.” (D’Arms and Jacobson 2006, 109)

The core relational theme of a given emotion type serves at least two purposes. First, as we have already seen in the case of guilt, the core relational theme spells out correctness conditions for emotions. Guilt is fitting if its object is in fact “guilt-worthy”. More generally, the question is whether a token emotional episode correctly presents a situation as falling under the core relational theme of the emotion type. This depends on whether the emotion correctly represents the situation as having the evaluative property picked out by the emotion’s core relational theme. Does, for example, the barking dog really constitute a threat? If not, my fear reaction to the dog’s barking is not fitting.

Second, the core relational theme also helps to distinguish one kind of emotion from another. Call this the Typologization Problem. What distinguishes fear from sadness is that the first homes in on dangers and threats, while the latter (roughly) is about what is lost. Since one is a future-directed attitude and the other backward-looking, they cannot be held with regard to the same situation or the same event. You can fear the bankruptcy of your firm or be sad about the bankruptcy of your firm, but you cannot both fear it and be sad about it at the same time because for you to fear the bankruptcy of your firm it cannot have happened yet, while for you to be sad about it must have occurred already.

It is important that these two uses of core relational themes are not independent from one another. This is made clear by the dispute concerning the adequate characterization of Sophie’s guilt rehearsed in the previous section. What we have seen there is that one way in which our emotional reactions can fail to be fitting is if we experience the wrong kind of emotion in a situation. Further, consider a possible gloss of the core relational theme of resentment: an affective response to actions manifesting ill will by others towards me. Given this gloss, it would be incorrect to react with resentment if someone treads on my hand by accident (Strawson, 1974). If the accident is due to negligence but not malevolence, the correct reaction would be anger (again, assuming a plausible core relational theme for anger). So the core relational theme for a given emotion spells out not only when that type of emotion is fitting, but also why *that* emotion is fitting rather than another.

To summarize, whether guilt is fitting depends on what the core relational theme of guilt is. But how do we determine core relational themes? So far I have only considered how the concept of core relational themes is used, but not *how* we arrive at the core relational theme of a given emotion type. I now turn to this question.

## **5. Sentimentalism and Core Relational Themes**

Emotion theories differ in terms of their metaethical commitments. How they propose to determine the respective core relational themes of emotion types will be shaped by these commitments.

Sentimentalists believe that values depend on emotions. The dependence is usually taken to imply both ontological and epistemic dependence: emotions both constitute value and detect value (D’Arms and Jacobson, 2006). Call this the Dependence Thesis. It follows from the Dependence Thesis that we do not have an independent, i.e., non-

affective access to the evaluative property that the emotion's core relational theme focuses on. Or if we do, that access is at best like the access the color-blind have to color-properties. For example, it is said in the case of fear that "the possession conditions of the concept of danger must mention the fundamental role of fear; and the property of being dangerous is one we can apply to objects *on the basis of fearing them.*" (Teroni 2007, 413) Or in the case of envy: "An episode of envy presents some difference in what might be termed 'position or possession' between the agent and some rival as being bad for the agent... just what counts as a possession, for the purposes of characterizing envy's concern, is driven not by some independent notion of the concepts *possession* or *rival...*" (D'Arms and Jacobson 2006, 106).

This metaethical view has important implications concerning how we arrive at – how we are to "distil" – the core relational theme of a given emotion type. The crucial claim is that we cannot fix the relevant evaluative properties in advance, but we must derive the evaluative property the core relational theme centers on by examining emotional responses themselves. This means that we can only individuate the property "\_\_\_is shameful", "\_\_\_is funny", or "\_\_\_is blameworthy" *via* the corresponding emotions of shame, amusement, or guilt. Accordingly, sentimentalists propose to establish core relational themes such as those sampled in the previous section by examining when and why people tend to respond with a certain kind of emotion, and what sort of actions that emotion motivates them to undertake. If we want to determine the core relational theme of envy, for example, we need to take a look at "the best interpretation of patterns in what people envy and why, and in the motivations people display when in the grip of that emotion." (D'Arms and Jacobson 2006, 106)

Of crucial importance for the characterization of patterns associated with a given emotion type are the typical eliciting conditions of the emotion and corresponding action tendencies. For example, it is argued that anger is typically elicited by perceived injustice or, more broadly, by violations of one's autonomy,<sup>15</sup> and typically issues in retributive actions or at least a desire to undertake such actions. Claims about typical elicitors and motivational effects can be further substantiated by studying the norms people invoke when justifying their emotional reactions,<sup>16</sup> and also enriched by evolutionary, phenomenological, and neurophysiological specifications.

What sentimentalists expect is that a sufficiently "thick" characterization of the pattern associated with a given emotion type will yield a reasonably convergent and univocal gloss of the emotion's core relational theme. With such a gloss in hand we will be able to say not just what evaluative property the core relational theme centers on, but also why the given evaluative property is such that it is *tailored to* or *calls for* this specific kind of emotion. We will be able to explain why that particular emotion was "set up to be set off" by that property (Prinz 2004, 65). Thus we will be able to say, for example, that perceived injustices or more broadly violations of one's autonomy call for anger because anger saliently represents to the agent that actions of this nature have been carried out against her, and at the same time anger motivates an effective response to such attacks (Prinz and Nichols, 2010). Or we can say that envy is a response tailored to the negative impact "the unmatched achievements of our rivals" can have on our well-being because envy makes this possibility salient for us, and through this very representation motivates us to undertake actions to offset this impact, e.g., by denigrating those achievements (D'Arms and Jacobson, 2006).

---

<sup>15</sup> Prinz and Nichols (2010, 125-30) consider both hypotheses and argue in favour of the broader one.

<sup>16</sup> For a cross-cultural study of the relationship between specific social norms and certain emotion types, namely emotions of interpersonal blame, see Rozin et al. (1999).



In the next section, I will try to assess how successful the sentimentalist method is in establishing core relational themes in the case of emotions central to the moral dilemma debate.

## 6. Sentimentalism and Moral Dilemmas

Can we use the conceptual device of core relational themes to determine whether Sophie is in a moral dilemma? The hope is that using the core relational theme of guilt we will be able to say whether guilt is a fitting response in Sophie's situation. Once the fittingness of guilt is established, friends of moral dilemmas could argue that Sophie is in a moral dilemma. Alternatively, using the core relational theme of regret we might reach the conclusion that in fact regret is the fitting reaction. Once the fittingness of regret is established, foes of moral dilemmas could claim that Sophie is *not* in a moral dilemma. I will argue that such hopes are bound to be disappointed. Invoking core relational themes does not help us make progress in the debate about moral dilemmas.

What then is the core relational theme of guilt? One sentimentalist suggestion for the core relational theme of guilt links guilt directly to moral responsibility and blameworthiness.<sup>17</sup> The suggested gloss is this: "someone I am concerned about has been harmed and I have responsibility for that in virtue of what I have done or failed to do." (Prinz and Nichols 2010, 134).

Now, if that is the right gloss of the core relational theme of guilt, then guilt is not a fitting reaction in Sophie's situation, and Sophie is not in a moral dilemma (as far as we can tell using the emotion-based approach). Prinz and Nichols make it clear that they mean here moral responsibility for harm voluntarily caused by the agent: "we feel guilty in cases where we think we could have prevented a bad outcome... where we [...] think we could have done otherwise" (Prinz and Nichols 2010, 135). But Sophie could not in fact have done otherwise in the sense that she did not have access to a morally better course of action than that involving the death of at least one of her children. If Sophie had not chosen one of her children to survive, more harm would have ensued as two children would have been killed, not one. Moreover, the one child who is killed, whether or not Sophie makes a choice, suffers the same harm. So the harm suffered by Sophie's child ensues not because of what Sophie has done or failed to do.<sup>18</sup>

If this core relational theme is adopted, and Sophie nevertheless does happen to feel guilt, then her guilt is best described as quasi-guilt akin to "survivor guilt", "authority

---

<sup>17</sup> Recall that according to the definition proposed at the outset of the paper the agent is blameworthy for an action if she is morally responsible for it and the action is unjustifiable.

<sup>18</sup> Some would disagree. It can be argued, for example, that what Sophie is blameworthy for is having *increased* the harm done to the child killed. The harm, it is said, is made worse because the child is sent to her death *by her own mother* (see Greenspan, 1983 for a discussion of this idea). Maybe so, but this does not mean that Sophie could have done better morally speaking. Perhaps the child does suffer additional harm due to the fact that her mother was involved in some sense in her death. However, had Sophie refused to choose, both children would have to suffer a similar (and possibly worse) kind of extra harm. Each can think: my mother was involved in my death – she could have prevented it, but did not. Again, the upshot is that Sophie does not have access to a morally better option. I would suggest the same response to those who think Sophie is blameworthy for having cooperated with evil if she accepts the offer. Perhaps by accepting the offer she does cooperate with evil and this could well be a bad thing (see esp. Kamm, 1999 on cooperation with evil). But this does not mean that she has a moral superior alternative course of action available to her (see Greenspan 1983, 122 for the same point on cooperating with evil). Now, it is a different question whether guilt can be fitting *even though* the agent has no access to a morally better option. This is a question I will address in the following. The point here is only that *if* Sophie's guilt is fitting it is not because she had access to a morally superior course of action and failed to choose it.

guilt”, and similar phenomena (Rawls 1971, 482). These are cases in which the agent feels guilt even though she is not in fact blameworthy because she has not done something unjustifiable. In such cases, the agent’s guilt is explicable, and perhaps even rational, but it is not fitting.<sup>19</sup>

However, others propose alternative core relational themes for guilt in terms of which guilt is fitting in Sophie’s situation. For example, D’Arms and Jacobson (1994, 743) gloss the core relational theme of guilt as “a negative emotion directed at one’s own action,” and even Prinz and Nichols consider this alternative core relational theme for guilt: “I have harmed someone whose well-being is a matter of concern to me”.<sup>20</sup> Indeed, in their analysis of a case structurally similar to *Sophie’s Choice*, D’Arms and Jacobson argue that guilt is the fitting emotional response as the relevant action has “violated the strongest wish of someone utterly dependent upon [oneself]” and has “sacrific[ed] [someone else’s] interests for the sake of [one’s] family.” But if guilt is fitting, then we must conclude that Sophie has done wrong and is blameworthy for doing so. In that case, however, we would also have to accept that Sophie is in a moral dilemma because guilt characterized in this way is fitting whatever she chooses.<sup>21</sup>

Now, which proposal should we accept? Should we accept a core relational of guilt as not requiring access to alternative courses of action that are morally better? Or are we instead to regard guilt experienced in Sophie’s situation as unfitting, and thus an “over-extension” of guilt’s core relational theme? The claim I want to make here is *not* that there is no answer to this question. Rather, my point is that sentimentalism simply does not have the resources to answer it. Recall that the sentimentalist proposal was to determine the core relational theme of an emotion by producing “the best interpretation” of eliciting conditions, characteristic phenomenology, action tendencies, normative social contexts, evolutionary history, and neurophysiological processes associated with a given emotion. I submit that there is simply no saliently best interpretation of the behavioral, phenomenological, neurophysiological patterns of guilt that would settle which of two core relational themes mentioned above we should embrace.

The fundamental difficulty is that how we should circumscribe the core relational theme of guilt crucially depends on cases such as *Sophie’s Choice*. The question that that situation raises is whether guilt can be fitting even though the agent had no alternative course of action available to her that would have been morally better than what she has actually done. My point here is that it is hard to see how the sentimentalist analysis of guilt will yield an answer to this question. Scrutinizing eliciting conditions, characteristic phenomenology, action tendencies, normative social contexts, evolutionary history, and neurophysiological processes associated with guilt will yield

---

<sup>19</sup> Prinz and Nichols (2010) claim that survivor guilt involves an “over-extension” of guilt’s core relational theme (134). However, they fail to explain why instead of modifying the core relational theme of guilt to accommodate cases such as “survivor guilt” we should instead rule out “survivor guilt” as not a case of guilt proper.

<sup>20</sup> Prinz and Nichols (2010, 133-4) eventually reject this suggestion for the core relational theme of guilt without really explaining what is wrong with it.

<sup>21</sup> As we will see shortly, D’Arms and Jacobson attempt to resist this conclusion. They claim that “guilt sometimes fits even when one has not acted wrongly” (2000a, 88n34). The first difficulty with this suggestion concerns the correctness conditions of guilt. Given the vagueness of “sometimes” it remains unclear when guilt is fitting. Why is guilt in moral dilemmas fitting but not in cases of “survivor guilt” or “authority guilt”? The second issue concerns the Typologization Problem. If guilt is fitting even when one has not acted wrongly, then its hard to see how its core relational theme differs from regret. D’Arms and Jacobson are of course aware of the problem. However, as I will try to show below, the solution they propose cannot be reconciled with the sentimentalist approach.

no determinate reply as to whether guilt in Sophie's situation is a fitting response or rather an "over-extension" of the evaluative concern of guilt.

Of course, one can try to evaluate Sophie's situation by directly tackling conceptual issues about moral dilemmas without appealing to emotions. One may discuss the plausibility of the ought-implies-can principle or insist on the conceptual incoherence of *any* course of action being *all-things-considered* unjustifiable. One may also appeal to the *unfairness* of holding an agent responsible who had no better options available. In fact, despite being sentimentalists, this is exactly how D'Arms and Jacobson proceed (D'Arms and Jacobson 1994, 753-4). They argue that even though guilt is fitting in cases like Sophie's this does not entail that there are moral dilemmas. Their reason for denying moral dilemmas is that they accept the ought-implies-can principle. It follows from that principle that an action cannot be unjustifiable and blameworthy if no better option was available. What we see here is that the rejection of moral dilemmas is based on the attractiveness of the ought-implies-can principle, and not on the sentimentalist analysis of guilt. It is because D'Arms and Jacobson accept that general principle but also regard guilt as fitting in cases relevantly similar to *Sophie's Choice* that they propose to redefine guilt as sometimes fitting even when the agent is not blameworthy and the action not unjustifiable.

As noted already, their specific proposal (i) fails to clarify the correctness conditions of guilt, and (ii) does not solve the Typologization Problem, i.e., explain how we are to distinguish guilt from regret. However, the main point is that such moves were not supposed to be made by sentimentalists in the first place. Since sentimentalists claim that "evaluation, and in particular moral evaluation, is somehow grounded in human sentiment" (D'Arms and Jacobson 2000b, 722), we would expect the relevant emotional response to be the point we depart from, not the point we arrive at. This follows from the Dependence Thesis according to which that we do not have an independent, i.e., non-affective access to the evaluative property that the emotion's core relational theme focuses on.<sup>22</sup>

In sum, I am skeptical whether sentimentalists can establish the core relational theme of guilt without contradicting their own Dependence Thesis. The appeal to emotions does not appear to do real work in the sentimentalists' thinking about moral dilemmas. Sentimentalists are also forced to determine the appropriate assessment of the relevant kind of situation independently from the corresponding emotional reactions. This belies the original promise of sentimentalism that we could move from emotional reactions to an identification of evaluative properties that those emotions, when fitting, are supposed to latch onto.<sup>23</sup>

Should sentimentalists be worried about this finding? I think they should for two reasons. The first problem is limited to the debate about moral dilemmas, the second

---

<sup>22</sup> It is customary to distinguish between subjective or simple sentimentalism (associated with the work of Haidt, Nichols, and Prinz, among others), on the one hand, and rationalist sentimentalism (D'Arms and Jacobson), on the other. The former is subjectivist and relativist (Haidt, 2001; Prinz, 2006; Prinz, 2007, etc.), while the latter allows for an assessment of the reasonableness of emotions in terms of non-subjective standards. For present purposes the differences between these two forms of sentimentalism can be ignored because they both accept the Dependence Thesis. See below on circularist versions of sentimentalism.

<sup>23</sup> It is for the same reasons why an ideal observer account would not be of much help here. First, the problem is not the lack of information (see footnote 2, also Greenspan, 1983). Second, we have seen that what is "normal" to feel is one thing, whether the emotion is fitting is another. Third, suppose the ideal observer would endorse guilt as fitting in Sophie's case. But why would an *ideal* observer do so? Perhaps because the ideal observer being in a "calm frame of mind" (Brandt 1959, 173) is capable of accessing the relevant evaluative property by non-affective means. That is indeed a possibility we cannot exclude. But then we have already moved past sentimentalism.

problem is more general. So the first problem is this. It seems that sentimentalists left to their own resources cannot resolve an important indeterminacy concerning the core relational theme of guilt. This is a crucial indeterminacy as far as emotion-based dilemma arguments are concerned. As long as it remains indeterminate whether or not the core relational theme of guilt implies that guilt can only be fitting if the agent could have done better morally speaking, we cannot use the core relational theme of guilt to establish whether guilt is a fitting response in Sophie's situation.

Sentimentalists may object at this point that this criticism is unfair because it starts at the wrong end. We should start instead by establishing why emotions taking center stage in emotion-based arguments are fitting in ordinary, paradigmatic cases. Based on such cases, those advocating the emotion-based pro-dilemma argument can obtain a general criterion for the fittingness of ordinary guilt. With this criterion in hand, they can then proceed to showing that guilt is the fitting reaction by Sophie. Those favoring the emotion-based anti-dilemma argument can do the same for regret.<sup>24</sup>

It is of course easy to find cases where the fittingness of guilt is not in question, or at least is far less contentious than in *Sophie's Choice*. To take a celebrated example (suggested by a reviewer), we would presumably all agree that if guilt ever is fitting, Raskolnikov's guilt in *Crime and Punishment* is a fitting response to the murder he committed. What we need to do, the sentimentalist could say, is to use such easy cases to fix the core relational theme of guilt, and then move onto more difficult cases involving various complications such as the lack of morally better alternatives.

But the question is not whether Raskolnikov's guilt is fitting. It clearly is a paradigm case of fitting guilt. The difficulty is what evaluative property renders it fitting. Again, my claim is not that there is no answer to this question. Rather, the problem is that it seems impossible to determine *with the required accuracy* what this evaluative property might be just by scrutinizing the emotional reaction itself and without drawing on independent evaluative considerations. Of course, we can relatively easily produce a convincing evaluation of what exactly is wrong with Raskolnikov's action once we draw on such independent considerations. And once the salient evaluative property of Raskolnikov's action and of other actions relevantly similar to his has been individuated, it will be possible to use the evaluative property we have narrowed in on as a criterion for the fittingness of guilt. However, this method is not the one recommended by the sentimentalist. The Dependence Thesis prescribes that we should scrutinize emotional reactions and derive the evaluative property on that basis. On this approach, we cannot individuate the evaluative property first and then use it as a corrective to the over-extension of the core relation theme.<sup>25</sup>

---

<sup>24</sup> This sentimentalist objection can draw support from an influential theory on what it takes to use concepts competently (Putnam, 1975). Competence depends on how the subject fares on central, stereotypical cases, e.g., whether she associates "is a tiger" with something like "large, orange, striped catlike creature". By extension, one can be said to be a sufficiently competent user of terms such as "guilt" and "blameworthiness" if one recognizes that it applies to central cases such as, say, Raskolnikov's (see below). One can leave it to experts to argue about cases like *Sophie's Choice*. For a negative assessment of sentimentalism that uses this approach, see Schroeter, 2006. I thank an anonymous reviewer for pressing me on this issue.

<sup>25</sup> Why not go circular? Perhaps we could obtain a rough-and-ready core relational theme using the various empirical methods mentioned earlier. With this core relational theme in hand, we might also be able to obtain some initial idea of the evaluative property the theme tracks. At this point, we could perhaps – through a dynamic and dialectical process of mutual adjustment – use the evaluative property to fine-tune the core relational theme, and conversely, use the core relational theme to more precisely circumscribe the property (see esp. Wiggins, 1987). Wiggins (also McDowell, 1998) argue that the circularity of this approach is not vicious. That may be so. However, the question is whether we have a sufficient initial grasp of the core relational theme enabling us to home in on the relevant evaluative property. I argue that at least in the case

I do not mean to deny that studying biological, psychological, and sociological features of a given emotion will reveal a great deal about the central evaluative concern of that emotion type. Doing so will no doubt “put us in the right ballpark” as regards the range of evaluative properties a given type of emotion could be said to track. The criticism I have been pressing here is only that this sort analysis will not produce a gloss of the core relational theme of the emotion that would be sufficiently fine-grained to do the work sentimentalists expect it to do. Specifically, the core relational theme of guilt cannot be defined using the sentimentalist method with a degree of precision that would enable us to adjudicate the debate between emotion-based arguments *pro* and *contra* the existence of moral dilemmas.

In the next section, I will discuss why the sentimentalist cannot be cavalier about this criticism.

## 7. Implications for Sentimentalism

So assuming that my interim conclusions above are correct, how damaging are they to the sentimentalist metaethical project? There is an obvious reason to worry if you are a sentimentalist and subscribe to an emotion-based *pro* or *contra* argument in the moral dilemmas debate. But many sentimentalists may be prepared to grant that there is not much mileage to emotion-based *pro* or *contra* dilemma arguments within a sentimentalist framework. At the same time, they can still reject the claim that this would have serious implications for sentimentalism. Let me address some considerations that the sentimentalist could invoke in support of this response.

A common reaction to the debate about moral dilemmas is that these are *recherché* cases of no great practical or even theoretical importance (e.g., Sunstein 2005, 541). The sentimentalist may reply in this vein too by arguing that “tragic cases” such as *Sophie’s Choice* are “rare and particular situations” (D’Arms and Jacobson 1994, 754).<sup>26</sup> Furthermore, the sentimentalist could insist that emotions such as guilt are *natural* psychological kinds (D’Arms and Jacobson, 2010). Consequently, it does not detract from the plausibility of sentimentalism that we will not be able to settle the question of the fittingness of emotional reactions in situations that are unlikely to ever occur in natural environments. Less radically, sentimentalists could allow that cases like *Sophie’s Choice* can be used to expose the vagueness of our evaluative concepts *at their limits*. Perhaps we should not expect that there would always be a determinate answer whether an action is blameworthy, a person cruel, a situation threatening, etc. Indeed, if we are realists about vagueness (see Pettit, 1991) the fact that our emotional responses can shade into one another may speak in favour of sentimentalism rather than against it.

I believe that the sentimentalist would be wrong to base her defense on the supposed rarity of dilemmatic situations. Many ethical theories accept or are at least compatible with the existence of moral dilemmas. For example, if there is a plurality of heterogeneous, incommensurable values, then moral dilemmas could be frequent.<sup>27</sup>

---

of guilt and its cognate emotions we do not. To repeat, the problem exposed by *Sophie’s Choice* is not a marginal issue. It diagnoses a central ambiguity regarding the identity of the relevant evaluative property.

<sup>26</sup> Though, somewhat oddly, a page earlier in the same text they say that “tragic cases are just one place where the problem we have been developing [i.e., the problem about warranted guilt] arises” (753).

<sup>27</sup> For example, Sartre’s story of a young Frenchman torn between the duty to defend his homeland and the duty to take care of his mother could be construed as a moral dilemma generated by the clash of incommensurable values (see McConnell, 1978).

In addition, there are many other examples that lack the special dilemmatic structure of *Sophie's Choice*, but pose the same kind of challenge to the sentimentalist analysis. Consider, for example, the case of the *Faultless Truck Driver* who runs over a child through absolutely no fault of his (Williams, 1976). Williams claims that this driver will feel a distinct kind of emotion. He calls this type of emotion "agent-regret". He then goes on to suggest that we can infer from this emotional reaction pattern that there is a distinct evaluative property that "agent-regret" latches onto.

In response, we can run the same argument regarding "agent-regret" as for *Sophie's Choice*: it is hard to see how we could come up with a sufficiently fine-grained core relational theme for agent-regret using the sentimentalist approach. Again, we can argue about whether the agent is morally answerable for some non-intended causal consequence of her intentional action. But studying the phenomenology, behavioral expressions, etc. of typical emotional reactions to the non-intended (harmful) effects of one's own intentional actions can hardly settle this argument.

But is it not sufficient to be in the right ballpark? Why would it undercut the sentimentalist project that it cannot deliver fine-tuned accounts of the core relational themes of emotions? It seems after all that by focusing on the empirical features of guilt we can work out a serviceable gloss of its core relational theme. For example, we can understand anthropologists talk about shame-cultures and guilt-cultures even if we do not have access to a precise definition of the core relational theme of shame or guilt.<sup>28</sup>

It may well be true that in certain contexts even coarse-grained characterizations of emotion types suffice. However, it is crucial to note that the sentimentalist's Dependence Thesis assigns a much more onerous task to emotions. What the Dependence Thesis implies is that when it comes to ascribing evaluative properties emotions are our highest court of appeal. They are meant to be the source of *ultima facie* justification (Tappolet, 2000). Based on the above considerations we have reason to doubt that core relational themes of guilt produced using the sentimentalist approach can successfully perform this function. The sentimentalist analysis leaves too indeterminate the identity of the evaluative property, to which guilt is supposed to be a fitting reaction, for guilt to play this justificatory role.

So I think what *Sophie's Choice* can help us see is that a lot depends on identifying this property as determinately as we can. Although the case itself maybe contrived, the issue at the heart of *Sophie's Choice* cannot be conveniently set aside. The question that that situation raises is whether an unjustifiable course of action can be unavoidable, that is, whether we can be blameworthy for our action even if we had no better alternative available. There is a whole range of philosophical debates that depend on this issue. These include most topics in the free will literature as well as the polemics about moral luck and involuntary sins, theological problems (e.g., theodicy), the problem of dirty hands in political philosophy, and many questions of distributive justice as well.

---

<sup>28</sup> Although even this might concede too much to sentimentalists. Empirical studies have found that boundaries between shame and guilt are often not at all pronounced in practice (Yik, 2010; Casimir and Schnegg, 2003). In one study, many of the subjects, in either the US or China, did not really distinguish between shame and guilt, either in terms of the subjective experience or in terms of concomitant appraisals and behaviour, i.e., "degree of moral stand, sense of responsibility and motivation to make amends". It is surprising that the Typologization Problem surfaces even for such basic emotions. It is surprising, and this again should worry sentimentalists especially, also because in much of contemporary ethics the distinction between shame and guilt is taken to stand for a crucial moral difference: "Yet the contrasts between feelings of guilt and feelings of shame are so striking that it is helpful to note how they fit in with the distinctions made between different aspects of morality." (Rawls 1971, 484; see also Wallace 1994, 38-40).

So it is not a marginal worry at the fringes of an otherwise well-defined concept of guilt whether fitting guilt requires that the agent have access to morally better alternatives. Depending on one's answer to that question the evaluative concern of guilt will be fundamentally different. And depending on one's answer to that question the action-guiding and motivational implications of guilt feelings will be different as well. For example, many think that guilt plays a key role in motivating the agent "to do better next time".<sup>29</sup> However, if the view one takes of guilt is that it can be fitting even when the agent could not have done something morally better than what she has done, then guilt need not have this specific motivational implication. Thus in some situations it will not be incoherent for an agent to think her guilt fitting without this feeling committing her to act differently should the same situation arise again.

## 8. Conclusion

Given how much depends on a fine-grained definition of the core relational theme of guilt, *Sophie's Choice* can be used to diagnose a common weakness of the sentimentalist framework itself.<sup>30</sup> Admittedly, this paper has only focused on emotions central to the moral dilemma debate: guilt, regret, and some cognate emotions. So another way for the sentimentalist to contain the implications of the criticisms made in this paper could be to argue that there is something special about emotions of self-blame. For some reason, one could speculate, it is particularly difficult to pin down the core relational theme of emotions belonging to this family, while this difficulty is absent in the case of other emotions.

Responding to this strategy in full is beyond the scope of this paper. Still, I note in closing that the strategy is unlikely to succeed. One important reason why has to do with the meshing, i.e., interdependence, of self-blaming and other-blaming emotions – a phenomenon frequently discussed by sentimentalists (Gibbard, 1990; D'Arms and Jacobson, 1994; Prinz, 2007; Prinz, 2011). Furthermore, sentimentalists themselves seek to establish the core relational themes of different emotion types using a piecemeal approach. Consequently, those critical of the sentimentalist approach are also entitled to evaluating separately various sentimentalist attempts at defining the core relational themes of important emotion types.

However, going beyond the dispute concerning the characterization of particular emotions, we can use the specific case discussed in this paper to articulate a general challenge facing sentimentalism as well. This general challenge is whether sentimentalists can produce non-trivial glosses of core relational themes of emotion types that are sufficiently fine-grained so that they can latch onto the evaluative properties our moral, aesthetic and other normative debates revolve around – and whether sentimentalists can do so without violating their own Dependence Thesis. The problems this paper points out with regard to guilt and related emotions suggest that overall pessimism may be in order as to whether sentimentalism will be able to master this challenge. Such worries complement other reasons we find in the literature for insisting on the "limits of sentimentalism" (Schroeter, 2006) as a metaethical project.

---

<sup>29</sup> "When plagued by feelings of guilt, say, a person wishes to act properly in the future and strives to modify his conduct accordingly." (Rawls 1971, 483).

<sup>30</sup> In this paper, I have not discussed the question how *non-sentimentalist* metaethical approaches fare in lending support to emotion-based arguments in the moral dilemmas debate. Marcus (1980), for example, uses a non-sentimentalist but emotion-based argument as part of her defense of the reality of moral dilemmas.

## REFERENCES

- BARTELS, D. M., and PIZARRO, D. A. 2011, "The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas", *Cognition* 121, pp. 154-161
- BRANDT, R. 1959, *Ethical Theory*, Englewood Cliffs: Prentice Hall
- CASIMIR, M. J. and SCHNEGG, M. 2003), "Shame Across Cultures: The Evolution, Ontogeny and Function of a 'Moral Emotion'", in: H. Keller, Y.H. Poortinga and A. Schölmerich, eds., *Between Culture and Biology: Perspectives on Ontogenetic Development*, Cambridge: Cambridge University Press, pp. 270-300
- DAHL, N. 1996, "Morality, Moral Dilemmas, and Moral Requirements", in: H.E. Mason, ed., *Moral Dilemmas and Moral Theory*, New York: Oxford University Press, pp. 86-101
- D'ARMS, J. and JACOBSON, D. 1994, "Expressivism, Morality, and the Emotions", *Ethics* 104, pp. 739-763
- D'ARMS, J. and JACOBSON, D. 2000a, "The Moralistic Fallacy: On the 'Appropriateness' of Emotions", *Philosophy and Phenomenological Research* 61, pp. 65-90
- D'ARMS, J. and JACOBSON, D. 2000b, "Sentiment and Value", *Ethics* 110, pp. 722-748
- D'ARMS, J. and JACOBSON, D. 2003, "The Significance of Recalcitrant Emotion (or, Anti-Quasijudgmentalism)", in: A. Hatzimoysis, ed., *Philosophy and the Emotions*, Cambridge: Cambridge University Press, pp. 127-146
- D'ARMS, J. and JACOBSON, D. 2005, "Sensibility Theory and Projectivism", in: D. Copp, ed., *The Oxford Handbook of Ethical Theory*, Oxford: Oxford University Press, pp. 186-218
- D'ARMS, J. and JACOBSON, D. 2006, "Anthropocentric Constrains on Human Value", in: R. Shafer-Landau, ed., *Oxford Studies in Metaethics I*, Oxford: Clarendon Press, pp. 99-126
- D'ARMS, J. and JACOBSON, D. 2010, "Demystifying Sensibilities: Sentimental Values and the Instability of Affect", in: P. Goldie, ed., *The Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press, pp. 585-613
- DÖRING, S. 2010, "Why Be Emotional?" in: P. Goldie, ed., *The Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press, pp. 283-301
- FEINBERG, J. 1978, "Voluntary Euthanasia and the Inalienable Right to Life", *Philosophy and Public Affairs* 7, pp. 93-123
- GIBBARD, A. 1990, *Wise Choices, Apt Feelings: A Theory of Normative Judgement*, Oxford: Clarendon Press
- GOWANS, C. W. 1996, "Moral Theory, Moral Dilemmas, and Moral Responsibility", in: H.E. Mason, ed., *Moral Dilemmas and Moral Theory*, New York: Oxford University Press, pp. 199-215
- GREENE, J. D., SOMMERVILLE R. B., NYSTROM L.E., DARLEY J. M., and COHEN J. D. 2001, "An fMRI Investigation of Emotional Engagement in Moral Judgment", *Science* 293, pp. 2105-2108



- GREENSPAN, P. S. 1983, "Moral Dilemmas and Guilt", *Philosophical Studies* 43, pp. 117-125
- HADT, J. 2001, "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment", *Psychological Review* 108, pp. 814-834
- KAMM, F. M. 1999, "Responsibility and Collaboration", *Philosophy and Public Affairs* 28, pp. 169-204
- KOENIGS, M., YOUNG L., ADOLPHS R., TRANEL D., CUSHMAN F., HAUSER M., and DAMASIO A. 2007, "Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgement", *Nature* 446, pp. 908-911
- LAZARUS, R. S. 1984, "On the Primacy of Cognition", *American Psychology* 39, pp. 124-129
- LAZARUS, R. S. 1991, *Emotion and Adaptation*, New York: Oxford University Press
- MARCUS, R. B. 1980, "Moral Dilemmas and Consistency", in: C. W. Gowans, ed., *Moral Dilemmas*, Oxford: Oxford University Press, pp. 188-204
- MARCUS, Ruth B. (1996): "More about Moral Dilemmas", in: H. E. Mason, ed., *Moral Dilemmas and Moral Theory*, New York: Oxford University Press, pp. 23-35
- Mason, H. E. (ed.) 1996, *Moral Dilemmas and Moral Theory*, New York: Oxford University Press
- MCCONNELL, T. 1978, "Moral Dilemmas and Consistency in Ethics", *Canadian Journal of Philosophy* 8, pp. 269-287
- MCCONNELL, T. 1996, "Moral Residue and Dilemmas", in: H. E. Mason, ed., *Moral Dilemmas and Moral Theory*, New York: Oxford University Press, pp. 36-47
- MCDOWELL, J. 1998, "Projection and Truth in Ethics", in: J. McDowell, ed., *Mind, Value, and Reality*, Cambridge, MA: Harvard University Press, pp. 151-166
- MOTHERSILL, M. 1996, "The Moral Dilemmas Debate", in: H. E. Mason, ed., *Moral Dilemmas and Moral Theory*, New York: Oxford University Press, pp. 66-85
- MULLIGAN, K. 2009, "Emotions and Values", in: P. Goldie, ed., *The Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press, pp. 475-500
- PETTIT, P. 1991, "Realism and Response-Dependence", *Mind* 100, pp. 587-626
- PRINZ, J. 2004, *Gut Reactions. A Perceptual Theory of Emotions*, Oxford: Oxford University Press
- PRINZ, J. 2006, "The Emotional Basis of Moral Judgments", *Philosophical Explorations* 9, pp. 29-43
- PRINZ, J. 2007, *The Emotional Construction of Morals*, Oxford: Oxford University Press
- PRINZ, J. 2011, "Sentimentalism and Self-Directed Emotions", in: A. Konzelmann Ziv, et al., eds., *Self-Evaluation*, *Philosophical Studies Series* 116, pp. 135-153.

- PRINZ, J. and NICHOLS, S. 2010, "Moral Emotions", in: J.M. Doris and The Moral Psychology Research Group, eds., *The Moral Psychology Handbook*, Oxford: Oxford University Press, pp. 111-146
- PUTNAM, H. 1975, "The Meaning of Meaning", *Minnesota Studies in the Philosophy of Science* 7, pp. 131-193
- RABINOWICZ, W. and RØNNOW-RASMUSSEN, T. 2004, "The Strike of the Demon: On Fitting Pro-Attitudes and Value", *Ethics* 114, pp. 391-423
- RAWLS, J. 1971, *A Theory of Justice*, Oxford: Oxford University Press
- ROZIN, P., HAIDT J., IMADA S., and LOWERY, L. 1999, "The CAD Triad Hypothesis: A Mapping between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity)", *Journal of Personality and Social Psychology* 76, pp. 574-586
- SCHROETER, F. 2006, "The Limits of Sentimentalism", *Ethics* 116, pp. 337-361
- STRAWSON, P.F. 1974, "Freedom and Resentment", in: P. F. Strawson, *Freedom and Resentment*, London: Methuen, pp. 1-28
- SUNSTEIN, C. R. 2005, "Moral Heuristics", *Behavioral and Brain Sciences* 28, pp. 531-573
- TAPPOLET, C. 2000, *Émotions et Valeurs*, Paris: Presses Universitaires de France
- TERONI, F. 2007, "Emotions and Formal Objects", *dialectica* 61, pp. 395-415
- YIK, M. 2010, "How Unique is Chinese Emotion?", in: M. H. Bond, ed., *The Oxford Handbook of Chinese Psychology*, Oxford: Oxford University Press, pp. 205-217
- WALLACE, J. R. 1994, *Responsibility and the Moral Sentiments*, Harvard University Press, Cambridge, MA
- WIGGINS, D. 1987, "A Sensible Subjectivism?", in: D. Wiggins, *Needs, Values, Truth*, 3<sup>rd</sup> ed., Oxford: Clarendon Press, pp. 185-214
- WILLIAMS, B. 1973, "Ethical Consistency", in: B. Williams, *Problems of the Self*, Cambridge: Cambridge University Press, pp. 166-186
- WILLIAMS, B. 1976, "Moral Luck", in: B. Williams: *Moral Luck*, Cambridge: Cambridge University Press, pp. 20-39
- ZIMMERMAN, M. J. 1988, *An Essay on Moral Responsibility*, Totowa: Rowman and Littlefield