P.F. Strawson on Punishment and the Hypothesis of Symbolic Retribution

ARNOLD BURMS, STEFAAN E. CUYPERS, AND BENJAMIN DE MESEL

Abstract

Strawson's view on punishment has been either neglected or recoiled from in contemporary scholarship on 'Freedom and Resentment' (FR). Strawson's alleged retributivism has made his view suspect and troublesome. In this article, we first argue, against the mainstream, that the punishment passage is an indispensable part of the main argument in FR (section 1) and elucidate in what sense Strawson can be called 'a retributivist' (section 2). We then elaborate our own hypothesis of symbolic retribution to explain the continuum between moral reactive attitudes and punishment that Strawson only adumbrates (section 3). After this justification of the punitive response to wrongdoing, we compare and contrast our specific kind of retributivist hypothesis with other positions in the so-called 'new retributivism' (section 4). Our hypothesis differs from other subvarieties of expressive retributivism in putting centre stage the idea of punishment as taking up a reverential stance towards the victim.

1. The Place of Punishment in 'Freedom and Resentment'

In 'Responsibility and the Limits of Evil', Watson (1987, pp. 255–58) raises a suspicion about the interpenetration of responsibility and the retributive sentiments in the passage of 'Freedom and Resentment' (FR) where Strawson (1962) connects the reactive attitudes of indignation and disapprobation with punishment. According to Watson, the passage is troubling because 'if such attitudes involve retributive sentiments […], then skepticism about retribution is skepticism about responsibility' (p. 257). In response to this threat to responsibility, he suggests that 'the retributive sentiments can in principle be stripped away from holding responsible' (p. 258). Is Watson's suspicion justified and his suggestion plausible?[1] We do not think so.

[1] Watson's suspicion about Strawson's retributivism is, furthermore, motivated by a concern about free will. Since retributive, desert-involving responsibility standardly presupposes free will, scepticism about free will (in particular, based on historical considerations) would imply scepticism about responsibility. So, if Strawson accepts retributivism, then he cannot sidestep the free will debate, which is often taken to be one of his desiderata. Whether Strawson really tries to sidestep this debate is, however, questionable; see De Mesel (2022).

Here is the central part of the troubling passage:

> Indignation, disapprobation, like resentment, tend to inhibit or at least to limit our goodwill towards the object of these attitudes, tend to promote an at least partial and temporary withdrawal of goodwill; […]. The partial withdrawal of goodwill which *these* attitudes entail, the modification *they* entail of the general demand that another should, if possible, be spared suffering, is, […], the consequence of *continuing* to view him as a member of the moral community; only as one who has offended against its demands. So the preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment is all of a piece with this whole range of attitudes of which I have been speaking. (FR, p. 23)

Many theorists have found this passage problematic. Although Strawson himself neither uses the term 'retributive sentiments (or emotions)' nor the term 'the retributive reactive attitudes', he has been thought to have retributive inclinations. And given the disreputable status of retributivism, his view on punishment has been either neglected or recoiled from. Angela Smith, for example, describes Strawson's view as follows: 'As P.F. Strawson recognized, the moral responses that figure in our moral practices form a natural continuum, from positive and negative appraisals, to the reactive attitudes of admiration and resentment, to explicit verbal and behavioral expressions of gratitude and reproach' (Smith, 2015, pp. 120–21). She immediately distances herself from Strawson in a footnote: 'Though I disagree with Strawson that legal punishment and a "preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment" belong on this same continuum' (n. 48).

Theorists, even good and faithful Strawsonians, do not want to get embroiled in issues of punishment. They prefer to restrict themselves to non-punitive, outward expressions of blame.[2] However, on our view, the troubling passage on punishment is an integral part of

[2] See, for instance, McKenna (2012, pp. 127–72). Also, Russell (1990) holds that 'Strawson's remarks regarding the implications of his views for the problem of punishment are, […], both brief and obscure' (p. 554); Russell (2013) no longer says anything about punishment and dissociates himself even from 'the morality system' of blaming.

Strawson's fundamental line of argument in FR. We agree with Nicholas Sars (2022) that the basic structure of Strawson's argument is underappreciated in the contemporary scholarship. In particular, two pivotal elements are neglected: the real subject matter of the article (what Sars (2022, p. 3) calls 'Strawson's *Target Considerations*') and the distinction between personal and moral reactive attitudes.

Strawson delineates his focal topic for discussion, on whose justification the pessimists (libertarians) and the optimists (compatibilists) disagree, in section I of FR: 'the concepts of moral obligation and responsibility […], and the practices of *punishing* and blaming, of expressing moral condemnation and approval, […] the notions of moral guilt, of blame, of moral responsibility […]' (p. 1, our italics). After a long and winding argument he finally comes back to his central point of attention in section VI: 'The concepts we are concerned with are those of responsibility and guilt, qualified as "moral", on the one hand – together with that of membership of a moral community; of demand, indignation, disapprobation and condemnation, qualified as "moral", on the other – together with that of *punishment*' (FR, p. 23, our italics). So, punishment is part and parcel of the subject matter and cannot be set aside. Just as a reminder, in section II Strawson sketches the antagonism between a desert-based and an efficacy-based justification of these concepts and practices; and in section VI he proposes a reconciliation between the two parties about the subject matter in light of his overall argument: 'And now we can try to fill in the lacuna which the pessimist finds in the optimist's account of the concept of moral responsibility, and of the bases of moral condemnation and *punishment*; […]' (FR, p. 22, our italics). Obviously, the troubling passage cannot just be thought of as an afterthought on Strawson's part; rather, it should be taken as a central thought at the culmination of the whole argument.

What Sars (2022, pp. 1–2) calls 'the orthodox reading of F&R' has conflated two distinct classes of reactive attitudes: the *personal* class and the *moral* one. We concur with Sars that Strawson makes a crucial and substantial distinction between two fields of study: the principal field in section V and the neighbouring one in section III and IV. Strawson indicates that in order to deal with the antagonism between the optimists and pessimists he is going to make a detour in a neighbouring field: 'I want to speak, at least at first, of something else; […]. Perhaps something like the issue between optimists and pessimists arises in this *neighbouring field* too; and since this field is less crowded with disputants, the issue might here be easier to settle; and if it is settled here, then it might become easier to settle it in the *disputant-crowded field*' (FR, p. 5, our italics). In the centre of the less crowded, neighbouring field, covered in sections III and IV, stand *resentment* and the other *personal* reactive attitudes, such as gratitude, forgiveness, love and hurt feelings, discussed in relation to excuses, the objective attitude and the thesis of determinism. In the centre of the disputant-crowded, principal field, entered into again in section V, stand *moral indignation* and the other *moral* reactive attitudes, such as moral disapprobation and condemnation, discussed once more in relation to excuses, the objective attitude and the thesis of determinism. So, Strawson treats the principal field *by analogy with* the neighbouring field: 'The reactive attitudes I have now to discuss [moral indignation, moral disapprobation] might be described as the sympathetic or vicarious or impersonal or disinterested or generalized analogues of the reactive attitudes I have already discussed [resentment, gratitude]' (FR, p. 15). Clearly, the distinction between the neighbouring and principal field of study parallels the distinction between the originally introduced personal reactive attitudes and their moral analogues. Strawson uses the qualifier 'moral' systematically in sections I–II and V–VI, but not at all in sections III–IV.

This key distinction is not just perspectival but substantial.[3] The difference between personal and moral reactive attitudes is not merely formal, but a function of the kind of demand (personal *versus* moral) involved: 'Don't you do that to *me*' in the personal case, 'You cannot do that to *others*' in the moral case. When I resent you, I am reacting *as an individual*. I react from the point of view of my personal interests and ideals, and I need not perceive the demand reflected in my personal attitude as a socially sanctioned one. In contrast, when I am indignant at you, I am reacting *as a member of a social group*. The demand reflected in my moral attitude is socially sanctioned, so others in the group can be expected to share it with me; if they share the demand, then they can normally express it towards you, in the same way as I did, through the reactive attitude of indignation.

One reason for the disavowal of punishment is that the orthodox reading of FR blurs the personal-moral distinction and, consequently, treats the class of reactive attitudes as a

[3] Among others, Hieronymi (2020, p. 8) interprets the distinction in a purely formal way, as merely a matter of perspective: 'In general, then, a reactive attitude is *x*'s reaction to *x*'s perception of or beliefs about the quality of *y*'s will toward *z*. In the impersonal reactive attitudes, *x*, *y*, and *z* are different persons. In the case of the personal reactive attitudes, the same person stands in for *x* and *z*. In the case of self-directed reactive attitudes, the same person stands in for *x* and *y*.'

whole. Of course, if one interprets this class indiscriminately, then it is hard to see how punishment can be on a par with 'such things as gratitude, resentment, forgiveness, love, and hurt feelings' (FR, p. 5). But Strawson specifically holds that there is a continuum only between *moral* reactive attitudes and punishment. To recoil from punishment because it does not mesh with the personal reactive attitudes is a category mistake. It is important to emphasize that the phrase 'this whole range of attitudes' at the end of the troubling passage does not refer to the whole class of reactive attitudes, but specifically to the class of *moral* reactive attitudes. Punishment is exclusively connected to moral responsibility, moral blame, indignation, and moral condemnation, that is, to a whole cluster of *moral* concepts.

Accepting Strawson's personal-moral distinction, one might admit that *moral* responsibility and *moral* blame are exclusively connected to the moral reactive attitudes, but still hold that punishment is a *legal* matter and as such only involves legal liability, not moral responsibility. Here it is important to appreciate Strawson's conception of morality, which he expounds in more detail in his 'Social Morality and Individual Ideal' (1961).[4] On this conception, the observance of a set of rules is a condition for the existence of a society. Rules, principles, and demands are moral in virtue of being socially sanctioned. Such a social morality bears, according to Strawson (1961, p. 48) a close relationship to *law*:

> […] I doubt if the nature of morality can be properly understood without some consideration of its relationship to law. It is not merely that the spheres of morality and law are largely overlapping, or that their demands largely coincide. It is also that in the way law functions to give cohesiveness to the most important of all social groupings we may find a coarse model of the way in which systems of moral demand function to give cohesiveness to social groupings in general.

Given this conception, moral responsibility and legal liability are closely connected and even intertwined categories.

As quoted above, Smith (2015, p. 121, n. 48) disagrees with Strawson that punishment belongs to a continuum to which resentment also belongs. But Strawson only emphasizes the continuity between punishment and the *moral* reactive attitudes of indignation and

[4] For our elaboration of the connection between FR and this earlier article, see De Mesel and Cuypers (2023).

disapprobation. Still, one might object that these attitudes are *not* continuous with the ‘infliction of suffering on the offender’ but only with *non-punitive blame*. Dealing adequately with the relationships between reactive attitudes, blaming, and punishing would require a separate paper. In response, we limit ourselves to three points. The first point is exegetical. Strawsonian accounts of non-punitive blame notwithstanding (Wallace, 1994; McKenna, 2012; Coates and Tognazzini, 2013), Strawson systematically connects blame to punishment: the moral subject is potentially ‘the subject of justified punishment, blame or moral condemnation’ (FR, p. 3). Second, it is questionable whether blame can be non-punitive and isolated from ‘hard treatment’. Blaming always seems to involve some social or psychological suffering of the blamee. If one were to appeal to a *purely inward* sort of blame with no outward expression or consequence, then our reply would be that this so-called ‘inward blame’ is conceptually derivative from outward blame and makes no sense apart from blaming practices. Third, blame is, of course, not the same as legal or state punishment. However, applying Strawson’s personal-moral distinction, we hold that *moral* blame, like indignation, presupposes socially sanctioned demands of a group. The same holds for punishment. Thus, at least in this respect, the moral reactive attitudes are continuous with punishment.

2. Is Strawson a Retributivist?

Against the backdrop of Strawson’s personal-moral distinction and his social conception of morality, we claim that punishment is an ineliminable aspect of his main argument in FR. Which view of punishment does Strawson then hold? His own description of punishment in FR is sketchy and fragmentary. Following the signposts he sets out, we try to determine the kind of theory of punishment he can be associated with.

In line with Strawson’s move towards reconciling the disputing parties one might expect that he holds a hybrid or mixed theory of punishment. Although the optimistic story shows an important lacuna, he assesses this story as the right one after a radical modification. There is nothing wrong with punitive practices for regulating behaviour and if these practices were inefficacious, then they should be modified or dropped altogether: ‘savage or civilized, we have some belief in the utility of practices of condemnation and punishment’ (FR, p. 24, cf. p. 27). However, the optimistic story exposes ‘a characteristically incomplete empiricism, a one-eyed utilitarianism’ (FR, p. 25), the lacuna in which can only be filled by the pessimist’s requirement: ‘the man who is the subject of justified punishment, blame or moral condemnation must really *deserve* it’ (FR, p. 3). To restore this vital thing, there is no need for an ‘obscure and panicky metaphysics’ (FR, p. 27) but only for an acknowledgment of ‘that complicated web of attitudes and feelings which form an essential part of the moral life as we know it’ (FR, p. 24). Although the radical modification of the optimistic story does not involve a resort to ‘contra-causal freedom’ (FR, p. 25) or another metaphysical formula, nothing less than a sense of desert, recovered from the facts as we know them, can fill in the lacuna. Strawson seems to hold that punishment can be justified by *attitude-based* desert: ‘Only by attending to this range of attitudes can we recover from the facts as we know them a sense of what we mean, i.e. of *all* we mean, when, speaking the language of morals, we speak of desert […] condemnation […]’ (FR, p. 24).[5] Without the inclusion of desert in the optimist’s story – admittedly, a *radical* modification – the story is unacceptable: ‘What *is* wrong is to forget that these practices, and their reception, the reactions to them, really *are* expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes’ (FR, p. 27).

In light of these signposts, Strawson’s theory of punishment seems to be desert-based and thus backward-looking. Given the standard classificatory criteria, such a theory is labelled ‘retributivist’, although Strawson himself does not use the label.[6] Contrary to expectations, we do not think he holds a mixed theory because he does not even give a partially utilitarian justification in terms of deterrence or reformation, notwithstanding his approving considerations about the efficacy of the punitive practices. Confronted with the optimist’s picture, Strawson (FR, p. 22) seems to share with the pessimist a sense of emotional shock:

[5] We agree with Alvarez’s (2021, p. 199) reminder: ‘One should remember that Strawson aims to incorporate the sense of desert that the pessimist is rightly shocked to find lacking in the optimist’s inadequate—because purely consequentialist—understanding of our practices of holding each other responsible […]’. Strawson’s rejection of a panicky libertarian-style desert to externally justify responsibility practices does not imply a repudiation of an attitude-based desert to internally justify them; it is precisely this internal desert he has on offer for the pessimist to fill the gap in the optimist’s story.

[6] For a discussion of these classificatory criteria in the 1960s, at the time of FR, see Honderich (1984; originally published in 1969). For a retribution theory, desert is a necessary condition of justified punishment.

> These practices are represented [by the optimist] solely as instruments of policy, as methods of individual treatment and social control. The pessimist recoils from this picture; and in his recoil there is, typically, an element of emotional shock. He is apt to say, among much else, that the humanity of the offender himself is offended by *this* picture of his condemnation and punishment.[7]

Strawson seems to subscribe to the classical retributivist idea that punishment is a kind of respect for – and even a right of – the offender as a rational agent within the moral community: 'The partial withdrawal of goodwill which *these* attitudes entail, the modification *they* entail of the general demand that another should, if possible, be spared suffering, is, rather, the consequence of *continuing* to view him as a member of the moral community; only as one who has offended against its demands' (FR, p. 23). Against this background, we work on the assumption that Strawson holds a retributivist theory *of some sort*.

Yet, to identify the moral reactive attitudes with 'retributive sentiments', as Watson (1987) does, or with 'retributive reactive attitudes', as Holmgren (2014) does, is tendentious. Given the disreputable status of the qualifier 'retributive', moral reactive attitudes then quickly become associated with 'vindictiveness or malice' (Watson, 1987, p. 258), anger, (blind) vengeance, and cruelty. We agree with Alvarez (2021, p. 198) that this association is off target: 'It is consistent with Strawson's picture that moral indignation […] should not

[7] In a footnote to FR, Strawson refers to the work of his former tutor Mabbott (1956), a retributivist. The passage from Mabbott's paper which Strawson most probably had in mind reads as follows: 'the belief that men can be cured of anti-social tendencies by punishment leads irresistibly towards "Brave New World" and "1984". What is *shocking* to most people about these Utopias […] is not the cruelty (for there need be none), nor the falsity of the creeds thus imposed, but the degradation and violation of human personality' (p. 308, our italics). Mabbott started the retributivist resistance against the then dominant utilitarian theories of punishment in his 1939 paper 'Punishment'. As for the optimist's picture, Strawson refers in footnotes to Nowell-Smith (1948) and Nowell-Smith (1954), the latter of which is a critical study of Campbell (1931), the pessimist (libertarian) who Strawson most probably had in mind. Nowell-Smith (1948) has two parts. Part one gives the then standard compatibilist answer to the free will problem (Moore, Hobart, Ayer) and says the problem is solved. Part two takes up the remaining problem as to moral responsibility, blame, and punishment: the problem of 'fittingness' or 'merit'. This problem about desert amounts to the problem of the justification of punishment. Mabbott 1956 (pp. 295–302) takes up this problem and explicitly discusses Nowell-Smith's utilitarian solution: 'Rewards and punishments […] are distributed not because certain actions directly "merit" them, but because some useful purpose is believed to be served by inflicting them' (1948, p. 56).

involve even an inclination or desire to punish, let alone vindictiveness or malice. What Strawson says is required is *acquiescence* to punishment […] since this acquiescence may in fact be intensely reluctant it is doubtful that, for Strawson, indignation *need* involve vindictiveness or malice—or related unsavoury feelings (sadism, masochism, etc.)'.

Three points are relevant in response to the objection from vengeance or cruelty. First, the moral reactive attitudes are not personal 'gut feelings' but reactions to the transgression of socially sanctioned demands: 'these attitudes of disapprobation and indignation are precisely the correlates of the moral demand in the case where the demand is felt to be disregarded. The making of the demand *is* the proneness to such attitudes' (FR, p. 23). Moral indignation is not self-standing but socially contextualized by demands, rules, and principles. Second, the association of moral indignation with 'a readiness to acquiesce in the infliction of suffering on an offender' takes place 'within the "institution" of punishment' (FR, p. 24). The punitive response is formally and strictly regulated by law. Taking both of these points together, the punitive response always takes place against a background of a social group or community, a legal system or state.[8] Third, punishment is not typically coloured by emotions boiling over, such as fits of anger and rage: 'I am not in the least suggesting that these readinesses to acquiesce, […], are always or commonly accompanied or preceded by indignant boilings […]' (FR, p. 24). Though the punitive response, as a continuation of moral indignation, can be cruel, it *need not* be. Strawson carefully and cautiously circumscribes punishment as a preparedness or readiness to *acquiesce* in the infliction of suffering on offenders. Such an acquiescence is not at all an active or eager attitude but only a passive acceptance or reluctant submission, and nothing more. In terms of Mackie's distinction, we might classify Strawson's view as *negative* or at most *permissive* but not as positive retributivism.[9]

[8] Here it is again important to appreciate Strawson's social conception of morality: 'A socially sanctioned demand is doubtless a demand made with the permission and approval of a society; and backed, in some form and degree, with its power' (1961, p. 38). Also, compare with the contrast between retributive punishment and vengeance in Nozick (1981, pp. 366–70). Revenge is personal, retribution is not. In contrast to vengeance, retributive punishment can be inflicted by someone with no personal tie to the victim. In addition, the imposer of retribution is committed to general principles mandating punishment in other similar circumstances.

[9] 'Within what can be broadly called a retributive theory of punishment, we should distinguish negative retributivism, the principle that one who is not guilty must not be punished, from positive retributivism, the principle that one who is guilty ought to be punished. We can indeed, add a third principle of permissive retributivism, that one who is guilty may be punished' (Mackie, 1982, p. 207).

Nevertheless, according to Strawson, the punitive response, no matter how intensely reluctant, is internally connected with moral indignation: ‘the preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment is *all of a piece with* this whole range of [moral reactive] attitudes […]’ (FR, p. 23, our italics). Punishment is not just an external addition to the moral reactive attitudes that also could be subtracted from them arbitrarily; rather, ‘we have here *a continuum of attitudes and feelings* to which these readinesses to acquiesce [in punishment] themselves belong’ (FR, p. 24, our italics). This is as far as Strawson’s theory of punishment goes, perhaps with the addendum that the continuum between moral reactive attitudes and punishment belongs to ‘the facts as we know them [which] supply an adequate basis for [our] concepts and practices’ (FR, p. 2). However, such a fact does not constitute an adequate *justification*. In FR Strawson does not explicitly address the problem of the justification of punishment. Yet all the elements of the problem are present. Strawson seems to acknowledge the possibility of justified punishment: ‘a readiness on the part of the offender to acquiesce in such infliction [of suffering] […] a readiness, […], to accept punishment as *“his due”* or as *“just”*’ (FR, p. 24, our italics). And he accepts the conundrum of holding that the ‘infliction of suffering on the offender […] is an essential part of punishment’, while complying with ‘the general demand that another should, if possible, be spared suffering’ (FR, p. 23). Yet, he leaves us in the dark as to *how* punishment might be justified or, at least, be permissible.

3. The Hypothesis of Symbolic Retribution

Starting from Strawson’s signposts, in particular the continuum between moral indignation and punishment and the reminder that practices which manifest our moral attitudes, including punitive practices, ‘do not merely exploit our natures, they express them’ (FR, p. 27), we propose our hypothesis of symbolic retribution as a constructive elaboration to deal with the ‘Strawsonian’ justificatory problem. We do not claim that our proposal factually underlies or fits best with Strawson’s own view on punishment in FR, which remains in any case unarticulated. We only offer our hypothesis as a plausible retributivist answer to the question of justification, which is at least compatible with Strawson’s pointers.[10] It is central to our

[10] As only one possible ‘Strawsonian’ hypothesis, our elaboration is not in competition with but complementary to, in particular, Bennett’s (2008, pp. 47–73) elaboration of Strawson’s argument as a retributivist ‘right to be punished’ strategy. Below, we come back to our kinship with Bennett.

hypothesis to realize that the structure of the punitive response to wrongdoing is similar to that of other expressive responses to dramatic events in our social and emotional life. The punitive response to crime is, therefore, a subclass of the class of these responses to tragedy. (Burms, 2005) Since legal punishment is relevantly analogous to certain nonlegal practices with which we are familiar, we introduce the punitive concept of symbolic retribution by giving first some *non-punitive* examples of expressive responses to devastating life events.

It is an anthropological datum that human beings feel the need, and have the desire, to respond to fatal events in life that cannot be made undone. After these devastating events occurred, mending and restoring their destructive effects are not possible anymore. Apart from reacting emotionally and taking steps, if possible, to prevent similar future occurrences, humans react in a special way to the unchangeable past. Because of the irreversibility of such serious events, humans typically also react *symbolically* to them.[11] Without doubt, honouring the dead is the most salient transcultural pattern of such a symbolic response. We give some examples.

Motivated by the need, or stronger, the obligation to be respectful toward the deceased we symbolically pay tribute to them by means of mourning and other funeral rituals. W.H. Auden (1936, p. 141) voices this symbolic tribute in his poem 'Funeral Blues', which begins as follows:

> Stop all the clocks, cut off the telephone,
> Prevent the dog from barking with a juicy bone,
> Silence the pianos and with muffled drum

[11] The word 'symbolic' has three dictionary meanings: (i) with the use of symbols (e.g., say it with flowers), (ii) symbolic of (e.g., the rose is a symbol of love) and (iii) symbolic, not literal. The last connotation is beset by a literal/real ambiguity in that it is ambiguous between (a) symbolic, not literal but real and (b) merely (or purely) symbolic, not literal and not real (e.g., the symbolic amount of one euro). The symbolic is always not literal, but the not-literal should not automatically be identified with the not-real. Only when the symbolic is qualified by 'merely (or purely)', it is not real in the sense that it has no, or little impact, that it is of no, or little consequence. Yet the unqualified symbolic has real impact, is of real consequence. A symbolic ritual, for example, is not literal but real because of its genuine effects on opinion, emotion, and conduct. We explicitly exclude the latter (sub)connotation (iii, b) from our use of 'symbolic'. Furthermore, on our use, symbols can be linguistic and non-linguistic, including objects (e.g., flag), states (e.g., silence), events (e.g., the last post), and processes (e.g., ritual).

> Bring out the coffin, let the mourners come.

One cannot go on with business as usual as if nothing happened. One must interrupt daily life by symbolically exhibiting the drama of death.

Similarly, something has to be done in honour of the victims of accidents, such as putting flowers at the place of the accident or observing a minute of respectful silence. The daily routines should be interrupted to symbolically pay tribute to the victims. Out of respect for the victims there would also be strong protests against the plan of transforming a former concentration camp into a supermarket. Such a place should symbolically remain untouched and kept clear of all commercial interference.

Relatedly, in a case of murder, Arnold Bennett (1908, pp. 252–56, our italics) describes in *The Old Wives' Tale* how the place of and the time around a murder get special significance demanding restraint, the neglect of which by ongoing mercantile routines and daily trivialities as if nothing happened evokes an emotional shock:

> The shop of the crime was closed, and the blinds drawn at the upper windows of the house. There was absolutely nothing to be seen, not even a policeman. Nevertheless the crowd stared with an extraordinary obstinate attentiveness at *the fatal building* in Boulton Terrace, […]. All had a peculiar feeling that the day was neither Sunday nor week-day, but *some eighth day of the week*. Yet in the St Luke's Covered Market close by, the stallkeepers were preparing their stalls just as though it were Saturday, just as though a Town Councillor had not murdered his wife […] he found customers, as absorbed in the trivialities of purchase as though nothing whatever had happened. He was shocked; he resented their callousness.

Against the backdrop of these examples, we hold, suitably hedged, that the punitive response to criminal offences is analogous to the non-punitive response to dramatic life events. As observance of a minute of silence counteracts the indifference of daily business towards fatal casualties, punishment counteracts the indifference of murderers towards the death of their victims. Punishment is a specific subclass of the general class of symbolic responses to tragedy. Hence, punishment is primarily a symbolic response.[12] Punishing wrongdoers is as symbolic as honouring the dead. The punitive response is more specifically an expressive response to past wrongdoing by wrongdoers who deserve such a response. Hence, punishment is symbolic retribution.

To forestall initial objections, our hypothesis of symbolic retribution should be qualified as follows. First, we restrict it to *serious* wrongdoing, more particularly, to the most serious moral as well as criminal offence of homicide, where 'the magnitude of the injury' (FR, p. 23) is maximal.[13] Second, we work under the idealized conditions that the offender has neither a justification nor an excuse for his wrongdoing, and is not exempted from liability either as, for instance, a psychopath might be. Since we assume *mens rea*, our hypothesis does not apply to crimes of strict liability. Third, the punitive response is internally justified, as Strawson indicates, on the basis of attitude-based desert, not on that of an external or metaphysical deservedness. Fourth, although the symbolic or expressive function of punishment is basic and intrinsic, it is compatible with at least some of the instrumental functions of punishment, as Strawson also indicates: 'savage or civilized, we have some belief in the utility of practices of condemnation and punishment' (FR, p. 24). Fifth, while Strawson 'confine[s] our attention to the case of the offenders' (FR, p. 23), it is crucial to our hypothesis to take into account the dynamics between (a) the offender, (b) the victim and, what we call, (c) the generalized other.[14]

[12] To underscore the symbolic character of the punitive response, we think it is instructive to refer to posthumous punishment (e.g., Melissaris, 2017) and animal punishment (e.g., Evans, 1906).

[13] Take, for instance, the Robert Harris case, as discussed by Watson (1987, pp. 235–42). We believe that it is possible to generalize our hypothesis to other serious crimes against a person (e.g., child abuse, rape) but we will not argue for this generalization here. We explicitly remain non-committal as for crimes against property (e.g., theft), statutory crimes (e.g., traffic offences), financial crimes (e.g., tax evasion), and victimless crimes (e.g., drug use).

[14] Moral indignation and disapprobation are 'the sympathetic or vicarious or impersonal or disinterested or *generalized* analogues of the [personal] reactive attitudes' (FR, p. 15, our italics). Following this terminology, we call *the third party* other than the victim and the offender 'the generalized other', who represents the community or the state. We borrow the term from Mead (1934, p. 154): 'The attitude of the generalized other is the attitude of the whole community'. In court, the independent judge and public prosecutor play officially and legally the role of the generalized other.

To see how our hypotheses of symbolic retribution explains the continuum between moral indignation and punishment consider the following dialectic, or better, 'trialectic' between victim, offender, and generalized other. The offender murdered the victim. The generalized other reacts with *moral indignation* to the offender and, in particular, his murderous act that killed the victim. Strawson limits himself to the relation between the generalized other and the offender. We also take into account the place of *the victim* in relation to both the offender and the generalized other. The offender's killing made an end to the one and only life of the victim. Since the offender irretrievably took the victim's life, the victim's death cannot be made undone. In this dramatic constellation, the relation of the generalized other to the victim is crucially important. The generalized other represents the moral community and embodies the moral demand. *We*, as a community, take ourselves seriously and cannot let the death of the victim pass. The generalized other cannot remain indifferent to what happened to the victim. The generalized other's moral demand and concomitant moral indignation – '[t]he making of the demand *is* the proneness to such attitudes [of disapprobation and indignation]' (FR, p. 23) – involve *taking up a reverential stance towards the victim*. We cannot shrug our shoulders when confronted with murder; we cannot go on as if nothing happened. Something has to be done out of respect for the victim. Action has to be taken as an honorary tribute to the victim who lost his only life. What has to be done, or what action has to be taken, cannot but involve *the offender* who caused the death of the victim.

We turn to the dialectic after the murder between the offender and the killed victim. Murder is the most serious offence and the harm done is absolutely irreparable. The offender cannot give the victim's life back; he cannot restore what he destroyed. The victim's life *was* unique, irreplaceable, and intrinsically valuable. Although the offender is in that sense in immeasurable 'debt' to the victim, he can never literally 'repay' the loss of the victim's life that he took. Nothing, literally nothing – no money, no goods, no services – can compensate for the irreparable harm done. Yet, something has to be done in honour of the victim, something which has to involve the offender. In taking up a reverential stance towards the victim, the moral community has to give some *public sign* directed towards the offender. Faced with the tragedy of the victim's death, the generalized other, who embodies the moral demand, demands in the name of the victim that the offender should take up his responsibility for what happened to the victim through his fault. The generalized other makes sure that the offender is brought to trial and that he recognizes, or is at least confronted with his offence. If guilty of the offence and condemned, then 'He'll pay for it' *in the symbolic sense by means of his punitive suffering*.

Since the offender cannot literally pay off his debt, he can only symbolically repay the loss of the victim's life. Lacking any literal means, there only remain symbolic means. Because the victim lost his only life and this loss is immeasurably weighty, the symbol of the repayment cannot but be life-involving and painful for the offender. The only thing the offender can do is to symbolically pay back the victim with *his being punished*. The symbol of the offender's repayment is his own suffering and repentance. This symbolic retribution by means of the punishment of the offender is the action that has to be taken in a case of absolutely irreparable harm done to the victim. The victim's life and death should symbolically be honoured. No other symbol than the state (or process) of being punished matches the seriousness of the offence. Nothing less than punishment can function as the vehicle of paying tribute to the victim.

The suffering of the offender as 'an essential part of punishment' (FR, p. 23) is not intended as suffering for its own sake – that would be cruel and vengeful – but as a reverential tribute to the victim. Imposing punitive measures is an *attempt* to direct the offender's attention to the symbolic rehabilitation of the victim. We hope that the offender himself will acknowledge that the victim's death is his fault and that he will remorsefully take up a reverential stance towards his victim. Although Strawson does not elaborate on the relation between offender and victim, he makes reference to the offender's *own* acceptance of 'his due' as part of the punitive process (FR, p. 24). It is important to observe that the content of punishment – reverence for the victim – is determined not so much by the intention of the judge or the expected uptake of the wrongdoer as by the *penal practice* of the community. On our Strawsonian hypothesis, the practice of punishment is continuous with the *moral* reactive attitudes. Moral reactive attitudes are reactions by a third party (in this case, the community) to the good or ill will manifested by a person towards another person (in this case, the ill will manifested by a wrongdoer towards a victim). Thus, to regard punishment as continuous with the moral reactive attitudes is to regard it as essentially involving three parties (wrongdoer, victim, community). Because existing expressivist accounts of punishment tend to focus on the wrongdoer and the community (see next section), there is a need for a complementary account which focuses on the victim and asks what punishment expresses towards the victim.

We argue that the punishing community acknowledges the loss of the victim’s unique life and important place in the community by expressing a reverential stance towards the victim.

Our hypothesis is only a rough sketch. In a fuller account, we have to deal with, at least, two related problems, in response to which we can only indicate the direction of our answers here.[15] The first problem concerns the proportionality of punishment to crime. We have restricted ourselves to the serious crime of homicide but do not at all commit ourselves to the death penalty or lifelong incarceration, let alone corporal punishment or physical suffering as a proportioned response. Taking punitive measures should neither be confused with inflicting carceral or physical violence nor with vengeance or cruelty. Yet, limiting the offender’s autonomy in some institutional way, which causes some social or psychological suffering, seems inescapable in taking serious crimes seriously. The second problem concerns the symbolic adequacy of the punitive response. We hold that some condemnation and ‘hard’, or not so hard treatment of the offender as a symbolic response to his murdering the victim is called for, because just saying ‘sorry’ to the victim’s next of kin, or putting flowers on the victim’s grave, or delivering community service is certainly not symbolically adequate.

As against these considerations, one might object that institutional punishment never could be adequate as an honorary tribute to, for instance, a victim with prison abolitionist convictions. In reply, we again apply Strawson’s personal-moral distinction. Punishment is not a personal response of an individual to another individual, but a *moral* response of a group or of an individual *as a member of a group* to an offending member of that group. So, to be adequate the punitive response need not express the values of the abolitionist victim but only those of the group. Holding a religious funeral for a rabid atheist would indeed show irreverence for his personal convictions, whereas punishing the abolitionist’s murderer or rapist, who violated socially sanctioned moral demands, still expresses the reverential stance towards the victim as a member of the community. Moreover, *if* one accepts (attitude-based) desert, the symbolic adequacy of punishment as an honorary tribute is less problematic, because the offender really deserves some suffering for destroying the victim’s life.

To sum up our hypothesis: we start from the thought that paying *punitive* honorary tribute to victims is structurally similar to paying non-punitive honorary tribute to casualties.

[15] For the first problem, see, e.g., von Hirsch (1996); for the second, see, e.g., Bennett (2008, pp. 33–36, 144–49). Strawson only touches on the first one: ‘their [indignation’s and disapprobation’s] strength is in general proportioned to what is felt to be the magnitude of the injury’ (FR, p. 23).

Within the constellation between offender, victim, and generalized other, the continuum of attitudes between moral indignation and punishment can then be justified in terms of symbolic retribution as follows. The generalized other’s moral demand and moral indignation require a reverential stance towards the killed victim. In light of this stance and the impossibility of literally giving back the victim’s life, the generalized other is *prepared to acquiesce in punishment* as the symbolic retribution of the offender to the victim. It is this indispensable symbolic repayment of the loss of the victim’s life that connects the generalized other’s moral indignation about the offender’s taking the only life of the victim with the generalized other’s willingness to accept punishment, be it reluctantly (an unwilling willingness), of the offender.

4. A New Retributivism

Although our hypothesis is not Strawson’s, it might be called ‘Strawsonian’ in that it is not only compatible with FR but also takes as its first premise the continuum between moral indignation, implying attitude-based desert, and punishment. Our own contribution consists in offering a justification for this continuum by making use of the idea of punishment as taking up a reverential stance out of respect for the victim. Here, we will neither distinguish our view from different strands in utilitarian thinking about punishment (deterrence, reform/rehabilitation, prevention/incapacitation) nor defend it against possible objections from this side. We limit ourselves to setting apart our hypothesis of symbolic retribution from other so-called ‘new retributivist’ theories. Limiting ourselves even more, we will neither discuss mixed theories that combine retributivist backward-looking with consequentialist forward-looking elements nor radically alternative theories (restorative justice, abolitionism).

Although there is no standard categorization of (new) retributive views of punishment, we think it is useful to minimally distinguish between (i) intrinsic, (ii) fairness, and (iii) expressive retributivism.[16] The intrinsic variety (still) holds that desert is metaphysical and suffering intrinsically good; the fairness variety maintains that the hardship of punishment

[16] For a more exhaustive categorization, see Boonin (2008, pp. 85–154, 171–180) and for a partly overlapping, less recent one, see Cottingham (1979). See, for the initial statement of (i) Davis (1972) and of (ii) Morris (1968). We are well aware of detailed criticisms of (sub)varieties of retributivism and wholesale rejections of retributivism at large. For the first type of objections, see e.g., Hanna (2009); for the second one, see e.g., Zimmerman (2011) and Caruso (2021). Here, we will not consider these disagreements.

brings back into balance the just distribution of benefits and burdens after the unfair advantage gained by crime. Our hypothesis belongs to the last category as a subvariety.

Expressive retributivism is indebted to Joel Feinberg's argument that, in contrast to other penalties, 'punishment is a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation, […]. Punishment, […], has a *symbolic significance* largely missing from other kinds of penalties' (Feinberg 1965, p. 98). However, Feinberg's characterization of 'the symbolic machinery of punishment' (p. 104) is multiply ambiguous.[17] He seems to conflate expression of moral emotions with that of moral judgements, and to treat expressive function on a par with symbolic significance. To disambiguate this we propose to distinguish between three subvarieties of expressive retributivism: (a) communicative expressivism, (b) emotive expressivism, and (c) symbolism. Another ambiguity that infects the machinery has to do with Feinberg's interpretation of the expressive/symbolic function not as a justification but as a defining feature of punishment.[18] We take all three subvarieties, including our own hypothesis, as offering a justificatory basis for punishment.[19]

According to communicative expressivism, punishment should not (or not only) control or manipulate the offender but first and foremost communicate a message to him as a moral and rational agent. In such a forceful communication, judgements of disapproval also take the form of hard treatment imposed on the offender as a consequence of his transgression of the law or his disconnection from the correct values of the community. The punitive communication can be unilateral or bilateral. Robert Nozick (1981, pp. 370–71) holds the first version: 'Retributive punishment is an act of communicative behavior. […]. The (Gricean) message is: this is how wrong what you did was. […] we might see punishment as an attempt to demonstrate to the wrongdoer that his act was wrong, not only to mean the act is wrong but to *show* him [by matching measures, e.g. incarceration] its wrongness'. Antony Duff (2001, p. 79) holds the second, more inclusive and extensive version: 'Although some theorists talk of the 'expressive' purpose of punishment, we should rather talk of its communicative purpose:

[17] For an initial disambiguation, see Skillen (1980), and for a further elaboration, see Primoratz (1989).

[18] For this point, see Hart (2008, pp. 239, 263).

[19] We acknowledge that some versions of expressivism, particularly the communicative one, have also been developed in the literature on *non-punitive blame* and *praise* by, e.g., Macnamara (2015) and Telech (2021). Discussing the (in)consistency with our three punitive subvarieties is beyond the scope of this paper.

for communication involves, as expression need not, a *reciprocal* and *rational* engagement. […] it [communication] appeals to the other's reason and understanding – the response it seeks is one that is mediated by the other's rational grasp of its content'. Punishment should not only express to the offender the generalized other's condemnation or censure for his crime but also communicate to him the imposition of penance in the hope of his response in terms of repentance, reform, and reconciliation. Punishment as 'a species of secular penance' (p. 106) acknowledges the wrongdoing in officially censuring it and involves furthermore a two-way communicative process in which the wrongdoer is hopefully reasons-responsive to the message of hard treatment.[20] Although we are sympathetic to this communicative subvariety, we believe that it focusses too exclusively on the relationship between the generalized other (community or state) and the offender to the neglect of the pivotal role of the victim in the punitive response.

According to emotive expressivism, punishment expresses emotions in an institutionalized way. Two versions can be distinguished, according to whether 'calm' or 'violent' passions are being expressed. Jay Wallace (1994, pp. 68–69) holds the first version: 'What is essential to the harmful moral sanctions [infliction of suffering] […] is their function of expressing the emotions of resentment, indignation, and guilt; […]. In expressing these emotions […] we are not just venting feelings of anger and hatred, […]; we are demonstrating our commitment to certain moral standards, as regulative of social life.'[21] The chief

[20] Duff's censure-plus-sanctions view should be distinguished from von Hirsch's (1996) partly retributivist censure-*without*-sanctions view which he combines with a partly utilitarian view based on prudential reasons. According to Duff, the censorial message comprises both the trial and the condemnation. For Duff's more recent allegiance to retributivism in the new modality, see e.g., his 2011 chapter. Within this subvariety Wringe (2016) makes a further distinction between 'communicative accounts on which the principal audience of penal communication is the offender' and 'denunciatory accounts on which the principal audience is society as a whole' (p. 57).

[21] We classify Wallace as a 'calm' version because he subscribes to some sort of cognitive theory of the emotions: 'The special force of *judgments* of moral blame [denunciation, censure] can […] be understood as consisting in the expression of these reactive attitudes. […] the emotions in question are not arbitrary feelings of disapprobation and dislike; rather, they have propositional contents that are fixed by their connection to moral obligations that we accept' (pp. 75, 77). In a similar vein, Bennett (2014, pp. 4453–55) explores, without endorsing, the possibility of justifying retributive punishment on the basis of judgement-sensitive or cognitive emotions.

spokesman of the second version is Murphy (2003) who argues that the vindictive passions of the generalized other are not irrational in that they are reliable signs of the flouting of our common values by offenders. Out of self-respect and respect for the moral order vindictiveness and revenge-taking are justifiable as a response to wrongdoing. As anger is being expressed in aggressive behaviour, the emotions of indignation and resentment are being expressed within the confines of the law in punitive behaviour. In expounding his view Murphy writes: 'What Peter Strawson calls the "reactive attitude" of resentment, directed toward wrongs and those who do the wrongs, is a paradigm example of such [vindictive] emotional response' (2003, p. 19). In light of our argument for Strawson's personal-moral distinction (in section 1) and his intensely reluctant retributivism (in section 2) we cannot go along with either of these versions. We should not treat punishment as an expression of resentment (first as well as second version), as that would clash with the personal-moral distinction, and we should not see punishment as essentially vindictive (second version), as that would clash with Strawson's reluctant retributivism.[22] Moreover, both versions focus too exclusively on the (emotive) role of the generalized other in the punitive response.

Although it is misleading to construe Strawson's view as 'a paradigm example of' emotive expressivism, punishment, on his view, is indisputably 'all of a piece with' (FR, p. 23) *indignation*, which is the key *moral* emotion in our hypothesis. What then is the role of indignation in the punitive response? We already elucidated the specific character of this moral reactive attitude in our reply above to the objection from cruelty and took note of its connection to attitude-based desert. We now add that the generalized other's moral indignation is *not* itself the justificatory basis for punishment but only signals the offender's transgression of the moral 'basic demand', the demand which reveals our moral 'basic concern' for the victim.[23] We agree with Watson (2014) that this basic demand (and concern) to be treated with regard and good will has primacy relative to the reactive attitudes and

[22] In all fairness to Murphy, he later tempered his strong retributivist convictions and became a reluctant retributivist: 'And where do these second and third thoughts [about retributivism] leave me? They leave me as what I will call a "reluctant retributivist"' (Murphy, 2007, p. 16).

[23] Here we invoke Watson's (2014, p. 17) terminology: 'Strawson identifies two components of human sociality as crucial here. First, we care deeply (and 'for its own sake') about how people regard one another. Second, this concern manifests itself in a demand or expectation to be treated with regard and good will. Following Strawson, let's call these the *basic concern* and the *basic demand* respectively'.

feelings. Yet, contrary to Watson, we think that Strawson *himself* subscribes to this primacy of the basic demand. Strawson writes that 'the making of the demand *is* the proneness to such attitudes' (FR, p. 23), and although this suggests that there is a connection between the demand and the attitudes, 'proneness' clearly suggests that the demand can be made without actually experiencing indignation.

Thus, Strawson is *not* a paradigm example of emotive expressivism. The punitive response is not so much an emotional reaction as it is an action demanding moral regard for the victim. The more suitable construal of Strawson's view is that, as we suggested above, the generalized other's making of the moral demand – and the occasionally accompanying moral indignation – involve *taking up a reverential stance towards the victim out of concern for this victim*. On our construal, it is this basic concern for the victim that is the ultimate justificatory basis for punishing the offender.

Our Strawsonian hypothesis of symbolic retribution, together with Bennett's (2008) theory that conceives of punishment as an apology ritual, belong to symbolism. In the light of the general human capacity for symbolic sensitivity and disposition to symbolic action this subvariety draws attention to the connection between punitive responses to wrongdoing and other non-punitive responses to certain events in life.[24] For our purposes, we have concentrated on funeral and mourning rituals in response to dramatic life events. Yet, the life events calling for a symbolic non-punitive response might be less dramatic, such as an accident or insult; they also might be undramatically positive, such as doing someone an unselfish service. In such more mundane cases 'saying sorry' or 'showing gratitude' respectively seems to be called for. Bennett starts from the informal everyday apology to develop his apology ritual account of punishment, according to which 'the key organizing principle of punishment should be making the offender do the sort of thing – that is, engage in the sort of apologetic action – that he would do willingly and spontaneously were he to be properly sorry for his wrongdoing' (Bennett, 2016, p. 215). Punishment as an apology ritual is then the symbol that embodies the collective condemnation of the offender for his violation of criminal law. Bennett's view concentrates only on the relation between the generalized other

[24] Some adherents of restorative justice also acknowledge this fundamental anthropological fact, see e.g., Braithwaite (2000).

and the offender, whereas our hypothesis emphasizes also, and more insistently, the relation between the generalized other and the victim.[25]

Symbolism does not deny the communicative and emotive functions of retributive punishment. Our distinctive hypothesis complements these other subvarieties of expressivism, as well as Bennett's apology view. Our new retributivism differs from these other versions – and even from our 'apologetic friend' – in the way the victim is prioritized, thus highlighting an element that is absent (or obscured) in its expressivist alternatives. Whereas communicative retributivism as well as apologetism focuses more on the offender and emotivism more on the generalized other, our symbolism primarily focuses on *the victim* in the threefold punitive dialectic. To put it slightly otherwise, whereas emotive expressivism is basically concerned only with the community, and communicative expressivism as well as Bennett's symbolism are basically about the community-offender relation, our Strawsonian symbolic retributivism involves the community-offender-victim 'trialectic' – the community makes the offender suffer out of respect for the victim.

Acknowledgements

We are very much indebted to Maria Alvarez and Christopher Bennett for their helpful comments on an earlier version of this paper. We are also grateful to two anonymous *Philosophy*-referees for their valuable criticisms. Research on this paper was funded by KU Leuven Internal Funds (project C14/22/007) and the Research Foundation Flanders (FWO) (project G026923N).

References

[25] After Hampton abandoned her moral education theory of punishment, she (1991) defended the retributive idea that the punitive response symbolizes the reaffirmation of the victim's value, thereby annulling the offender's message of the victim's inferiority. Although Hampton's view is congenial to ours in focusing on the victim, it uses a more abstract axiological terminology and comes closer to Kantian/Hegelian moral balance theory or fairness retributivism.

Maria Alvarez, 'P.F. Strawson, Moral Theories and "The Problem of Blame": "Freedom and Resentment" Revisited', *Aristotelian Society Supplementary Volume*, XCV (2021), 183–203.

Wystan Hugh Auden, 'Funeral Blues' (1936), in W.H. Auden, *Collected Poems*, Edward Mendelson (ed.), (New York: Vintage International, 1991), 141.

Arnold Bennett, *The Old Wives' Tale* (1908), (London: Penguin Books, 2007).

Christopher Bennett, *The Apology Ritual. A Philosophical Theory of Punishment* (Cambridge: Cambridge University Press, 2008).

Christopher Bennett, 'Retributivist Theories', in Gerben Bruinsma and David Weisburd (eds.), *Encyclopedia of Criminology and Criminal Justice* (New York: Springer, 2014), 4446–56.

Christopher Bennett, 'Punishment as an Apology Ritual', in Chad Flanders and Zachary Hoskins (eds.), *The New Philosophy of Criminal Law* (London: Rowman and Littlefield, 2016), 213–30.

David Boonin, *The Problem of Punishment* (Cambridge: Cambridge University Press, 2008).

John Braithwaite, 'Repentance Rituals and Restorative Justice', *The Journal of Political Philosophy*, 8:1 (2000), 115–31.

Arnold Burms, 'Retributive Punishment and Symbolic Restoration: A Reply to Duff', in Erik Claes, René Foqué, and Tony Peters (eds.), *Punishment, Restorative Justice and the Morality of Law* (Antwerp/Oxford: Intersentia, 2005), 157–164.

Charles Arthur Campbell, *Scepticism and Construction*. *Bradley's Sceptical Principle as the Basis of Constructive Philosophy* (London: George Allen & Unwin, 1931).

Gregg D. Caruso, *Rejecting Retributivism. Free Will, Punishment, and Criminal Justice* (Cambridge: Cambridge University Press, 2021).

D. Justin Coates and Neal A. Tognazzini (eds.), *Blame. Its Nature and Norms* (Oxford: Oxford University Press, 2013).

John Cottingham, 'Varieties of Retribution', *The Philosophical Quarterly*, 29: July (1979), 238–46,

Lawrence H. Davis, 'They Deserve to Suffer', *Analysis*, 32:4 (1972), 136–40.

Benjamin De Mesel, 'Taking the Straight Path: P.F. Strawson's Later Work on Freedom and Responsibility', *Philosophers' Imprint*, 22:12 (2022).

Benjamin De Mesel and Stefaan E. Cuypers, 'Strawson's Account of Morality and its Implications for Central Themes in "Freedom and Resentment"', *The Philosophical Quarterly*, published online 2023.

R. Antony Duff, *Punishment, Communication, and Community* (Oxford: Oxford University Press, 2001).

R. Antony Duff, 'Retrieving Retributivism', in Mark D. White (ed.), *Retributivism* (Oxford: Oxford University Press, 2011), 3–24.

Edward Payson Evans, *The Criminal Prosecution and Capital Punishment of Animals* (1906), (London: Faber and Faber, 1987).

Joel Feinberg, 'The Expressive Function of Punishment', in Joel Feinberg, *Doing and Deserving* (Princeton, NJ: Princeton University Press), 95–118.

Herbert Lionel Aldophus Hart, *Punishment and Responsibility*, 2nd edition (Oxford: Oxford University Press, 2008).

Jean Hampton, 'A New Theory of Retribution', in R.G. Frey and Christopher W. Morris (eds.), *Liability and Responsibility* (Cambridge: Cambridge University Press, 1991), 377–414.

Nathan Hanna, 'The Passions of Punishment', *Pacific Philosophical Quarterly*, 90:2 (2009), 232–50.

Pamela Hieronymi, *Freedom, Resentment and the Metaphysics of Morals* (Princeton, NJ: Princeton University Press, 2020).

Margaret R. Holmgren, 'A Moral Assessment of Strawson's Retributive Reactive Attitudes', in David Shoemaker and Neal A. Tognazzini (eds.), *Oxford Studies in Agency and Responsibility, Volume 2* (Oxford: Oxford University Press, 2014), 165–86.

Ted Honderich, *Punishment. The Supposed Justifications* (Cambridge: Polity Press, 1984).

John David Mabbott, 'Punishment' (1939), in H.B. Acton (ed.), *The Philosophy of Punishment* (London: Macmillan, 1969), 39–54.

John David Mabbott, 'Freewill and Punishment', in H.D. Lewis (ed.), *Contemporary British Philosophy. Personal Statements* (London: George Allen & Unwin, 1956), 289–309.

John L. Mackie, 'Morality and the Retributive Emotions', in John L. Mackie, *Persons and Values* (Oxford: Clarendon Press, 1982), 206–19.

Coleen Macnamara, 'Blame, Communication, and Morally Responsible Agency', in Randolph Clarke, Michael McKenna and Angela M. Smith (eds.), *The Nature of Moral Responsibility* (Oxford: Oxford University Press, 2015), 211–35.

Michael McKenna, *Conversation and Responsibility* (Oxford: Oxford University Press, 2012).

George Herbert Mead, *Mind, Self, and Society from the Standpoint of a Social Behaviorist* (1934), Charles W. Morris (ed.), (Chicago: The University of Chicago Press, 1962).

Emmanuel Melissaris, 'Posthumous "Punishment": What May Be Done About Criminal Wrongs After the Wrongdoer's Death?', *Criminal Law and Philosophy* 11:2 (2017), 313–29.

Herbert Morris, 'Persons and Punishment', *Monist,* 52:4 (1968), 475–501.

Jeffrie Murphy, *Getting Even. Forgiveness and its Limits* (Oxford: Oxford University Press, 2003).

Jeffrie Murphy, 'Legal Moralism and Retribution Revisited', *Criminal Law and Philosophy* 1:1 (2007), 5–20.

Patrick Horace Nowell-Smith, 'Freewill and Moral Responsibility', *Mind*, LVII: Jan. (1948), 45–61.

Patrick Horace Nowell-Smith, 'Determinist and Libertarians', *Mind*, LXIII: July (1954), 317–37.

Robert Nozick, *Philosophical Explanations* (Cambridge, MA: The Belknap Press of Harvard University Press, 1981).

Igor Primoratz, 'Punishment as Language', *Philosophy*, 64: April (1989), 187–205.

Paul Russell, 'Hume on Responsibility and Punishment', *Canadian Journal of Philosophy,* 20:4 (1990), 539–63.

Paul Russell, 'Responsibility, Naturalism, and "The Morality System"', in David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility, Volume 1* (Oxford: Oxford University Press, 2013), 184–204.

Nicholas Sars, 'Strawson's Underappreciated Argumentative Structure', *European Journal of Philosophy,* early view 2022.

Anthony J. Skillen, 'How to Say Things with Walls', *Philosophy* 55: Oct. (1980), 509–23.

Angela M. Smith, 'Responsibility as Answerability', *Inquiry* 58:2 (2015), 99–126.

Peter F. Strawson, 'Social Morality and Individual Ideal' (1961), in Peter F. Strawson, *Freedom and Resentment and Other Essays* (London: Routledge, 2008), 29–49.

Peter F. Strawson, 'Freedom and Resentment' (1962), in Peter F. Strawson, *Freedom and Resentment and Other Essays* (London: Routledge, 2008), 1–28.

Daniel Telech, 'Praise as Moral Address', in David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility, Volume 7* (Oxford: Oxford University Press, 2021), 154–81.

Andrew von Hirsch, *Censure and Sanctions* (Oxford: Oxford University Press, 1996).

R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press, 1994).

Gary Watson, 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme' (1987), in Gary Watson, *Agency and Answerability* (Oxford: Clarendon Press, 2004), 219–259.

Gary Watson, 'Peter Strawson on Responsibility and Sociality', in David Shoemaker and Neil A. Tognazzini (eds.), *Oxford Studies in Agency and Responsibility, Volume 2* (Oxford: Oxford University Press, 2014), 15–32.

Bill Wringe, *An Expressive Theory of Punishment* (Basingstoke: Palgrave Macmillan, 2016).

Michael Zimmerman, *The Immorality of Punishment* (Peterborough: Broadview Press, 2011).

ARNOLD BURMS (arnold.burms@kuleuven.be) *is emeritus professor in philosophy at the Institute of Philosophy, University of Leuven–KU Leuven. His research interests include moral luck, personal identity, and moral taboos.*

STEFAAN E. CUYPERS (stefaan.cuypers@kuleuven.be) *is professor in philosophy at the Institute of Philosophy, University of Leuven–KU Leuven. His research interests include moral/legal responsibility, mind-body relation, and education.*

BENJAMIN DE MESEL (benjamin.demesel@kuleuven.be) *is assistant professor in moral philosophy at RIPPLE (Research in Political Philosophy and Ethics Leuven), Institute of Philosophy, University of Leuven–KU Leuven. His recent publications include* P.F. Strawson and his Philosophical Legacy *(co-edited with Sybren Heyndels and Audun Bengtson, Oxford*

University Press, 2023), Strawson’s Account of Morality and its Implications for Central Themes in ‘Freedom and Resentment’ *(with Stefaan E. Cuypers, The Philosophical Quarterly, 2024), and* Taking the Straight Path. P.F. Strawson’s Later Work on Freedom and Responsibility *(Philosophers’ Imprint, 2022).*