Fair play retributivism and the problem of punishment

Göran Duus-Otterström

1. Purpose and aims

The purpose of this project is to assess the strength of so called *fair play retributivism* (FPR) as a solution to the problem of punishment, i.e. the problem concerning on what grounds, if any, punishment of lawbreakers by the state is justified. FPR holds that punishment is permissible, and normally required, as a means of restoring the fairness between the lawbreaker and the law-abiding. The lawbreaker, this theory holds, is free riding on the social scheme of cooperation; she accepts the benefits of the scheme without accepting the behavioral restrictions upon which it depends. The lawbreaker thus takes unfair advantage over the law-abiding. Punishment is justified in that it removes that advantage.

FPR holds out a lot of promise as an approach to punishment, not least as it might overcome the problems associated with standard consequentialist and retributivist approaches to punishment. Yet it is confronted by a set of powerful objections. The project aims to assess the merits of FPR in light of these objections, and to arrive at a reasoned conclusion as to whether the theory is a plausible answer to the problem of punishment. Part of the work needed to meet this evaluative aim, however, is conceptual and constructive, for the fact is that FPR, while promising, is insufficiently fleshed out and understood. So the project aims to do two things:

1. To construct, based on previous accounts of FPR and relevant branches of political theory, the best (in the sense most plausible) statement of fair play retributivism.
2. To assess the strength of that theory as an answer to the problem of punishment by

way of analyzing a set of objections to it. Thus, the project belongs to the field of normative political theory. It seeks to evaluate, through conceptual and normative analysis, a promising if underappreciated theory of punishment. As becomes evident below, I am positively predisposed to FPR. Yet I am genuinely unsure as to whether it can sustain the critique to which it is subjected. A likely outcome of the project is that FPR provides an important part of the correct theory of punishment, but that it needs to be qualified and complemented in various ways.

2. Survey of the field

Punishment is the "imposition of something which is intended to be burdensome or painful, on a supposed offender for a supposed crime, by a person or body who claims to have authority to do so" (Duff 2008b). In its familiar *legal* form, punishment is the state's imposition of pain or deprivation (e.g. imprisonment) on a person for transgressing the laws of the jurisdiction. Thus understood, punishment is by its very nature controversial (although perhaps not subject to fundamental controversy).[1] It involves intentional harm to persons, and intentional harm is normally forbidden. Punishment thereby raises stringent demands on justification – we need to think hard about on what grounds, if any, such treatment is justified (Berman 2006; Boonin 2008). This is important not least considering how *widespread* the practice is.

[1] As Rawls (2001) once noted, surprisingly few believe that punishment is unjustifiable. For the contrary abolitionist view, see e.g. Boonin (2008) and Hanna (2009).

Göran Duus-Otterström är verksam som forskare och lärare vid Statsvetenskapliga institutionen, Göteborgs universitet. E-post: goran.duus-otterstrom@pol.gu.se

Punishment is endemic to our societies. For example, in 2009 courts in Sweden meted out 13 600 prison sentences. Furthermore, fines were the main punishment in 32 800 court rulings (Brå 2010). This staggering amount of interference with people's liberty and property adds up to what is probably the most significant and intrusive of state practices, at least in peace time. We have a duty to think seriously about this practice.

The "problem of punishment", as I use that term, refers to the problem concerning precisely on what grounds, if any, punishment is justified. It is a classic problem within political theory and moral philosophy that over the years has sparked considerable debate. The debate can be partitioned into three broad camps. *Consequentialist* theories hold that punishment is justified since it contributes to societies' overall utility or welfare, for example in virtue of deterring future crime or incarcerating dangerous individuals. *Retributivist* theories argue that punishment is justified, on backward-looking grounds, by being an instance of paying back to the offender what he or she deserves. Finally, there are *hybrid* theories that seek to combine consequentialist and retributivist views into a supposedly more appealing mix. Each camp is associated with significant problems. Consequentialism is criticized for in principle allowing whichever disproportionate or otherwise undeserved punishment that contributes to overall welfare. Retributivism, in supporting punishment on grounds separate from its consequences, is charged with supporting pointless or even counterproductive infliction of suffering. Finally, hybrid theories, while until recently very popular, have been criticized for either collapsing back into simple consequentialism or else suffering from incoherence. In short, the debate is in flux and no alternative enjoys anything even approximating universal consent among theorists (Duus-Otterström 2007).

Getting into a bit more detail, the rival of consequentialism is usually taken to be

intrinsic-good retributivism (Honderich 2006). This is the view that criminal wrongdoing deserves punishment, and that punishment, when in proportion to what the offender deserves, is intrinsically good or just (see e.g. Moore 1998). This view can explain many of the convictions most people have when they think about punishment: that punishment must be (likely to be) deserved if it is to be justified; that its severity must "fit" the crime; and that punishments should not be distributed with an eye to overall consequences. Yet intrinsic-good retributivism also suffers many problems. For example, in supporting even punishment that is good for no one, it seems to support the imposition of pain for pain's sake (Dolinko 1991, Shafer-Landau 1996). Many simply do not buy this way of looking at punishment, preferring instead to side with Bentham's classic remark that "all punishment is mischief: all punishment in itself is evil /.../ it ought only to be admitted as far as it promises to exclude some greater evil" (1988, p. 170). Still, consequentialist or hybrid theories seem unable to account for many of the desert-based convictions that most support when it comes to punishment, or at least unable to account for them *in the right way*.[1] Hence, scholars have sought to describe a form of retributivism that neither relies on the simple intrinsic goodness of deserved punishment, nor collapses into consequentialism.

In *Punishment and Personal Responsibility* (2007) I made such an attempt. I there outlined a contractualist argument to the effect that a retributive penal regime, when

---

1   Consequentialists often advocate retributive considerations indirectly, deeming them to be essential to the maximization of welfare. Bentham (1988) for example spent a great deal of time working out a conception of proportionality in the context of his wider utilitarian theory of punishment. Yet such moves are unconvincing to many as they do not seem to value justice or fairness in the right way. As Kymlicka (2002) notes, what is wrong about unfair treatment is not so much its disutility, but its unfairness.

compared to its alternatives, is reasonably preferable to all. This is so, I argued, for two reasons. On the one hand, a retributive penal regime seems better able to secure legal security, which is something which we all can and should value. On the other hand, retributive punishment, which is administered with an eye to holding the offender responsible in a fundamental, blame-entailing sense, treats offenders with more respect than sanctions that aim to produce deterrence or rehabilitation. I referred to these reasons, respectively, as the *institutional* and *symbolic* reason. While this analysis, I now believe, got some things right and pointed out values I am still committed to, it was also significantly flawed in other respects. What my argument justified was not so much *punishment* as a particular "penal regime", and while I argued that we can support such a regime on retributivist grounds, it can just as well be supported on consequentialist ones. Moreover, the analysis only argued the supremacy of a retributive penal regime when compared to two salient alternatives, namely one that seeks deterrence and one that seeks rehabilitation. I did not address the more fundamental question whether the retributive penal regime I favored is preferable by all *to no penal regime at all.* The theory which promises to right these wrongs, while keeping with the contractualist-retributivist outlook, is FPR.

FAIR PLAY RETRIBUTIVISM

FPR can be defined as the view that punishment is justified on account of it being required to promote *fairness* in a society. The starting point is that offenders violate the social contract that underpins society. They choose to violate the mutually advantageous rules that bind us all, thus enjoying the benefits of those rules without accepting their costs. In that sense they gain an unfair advantage over the law-abiding. Fairness then requires that the offender be punished, thus restoring the fundamental "legal equality" of society. This view is distinctively retributivist

in that it justifies punishment on backward-looking grounds. Punishment is justified simply because it is required to remove the already acquired unfair advantage of the offender. Moreover, proportionality is entailed by FPR as greater punishment is reserved for greater unfairness. Yet FPR is not dependent on the seemingly mysterious or vengeful "bite back" impulses that intrinsic-good retributivism seems to rely on (Mackie 1982). It is dependent on the value of fairness, the fundamental commitment to which is not mysterious. For example, we often support policies that improve equality of opportunity among children simply because it is *fair*, irrespective of further consequences. If punishment could be portrayed as a matter of fairness, much would be gained in terms of retributivism's plausibility. This is the promise held out by FPR.

There is by now a decent-sized literature on FPR. The origins of the broad idea of thinking about (retributive) punishment in terms of fairness is usually credited the seminal papers of Morris (1968) and Murphy (1973). The approach was crystallized by Dagger (1993). It is now subject to growing attention, attracting defenders (Davis 1993, Dagger 2008, Hoskins 2010, Stichter 2010) as well as critics (Dolinko 1991, Boonin 2008, Duff 2008b).

FPR fits nicely with many of the standard theories in liberal political theory. Rawls famously characterized society as a "cooperative venture for mutual advantage" which "makes possible a better life for all than any would have if each were to live solely by his own efforts" (Rawls 1999, p. 4). The terms of social cooperation, however, is undergirded by binding rules of conduct, and it is essential to the functioning of society that those rules are recognized as binding by all. It is from the benefits of social cooperation that the basis of each citizen's obligation to follow the rules is derived. Rawls and others argue that we have *obligations of fair play*: to honor the terms of cooperation from which we benefit

by respecting those very terms. As Hart put the point: "when a number of persons conduct any joint enterprise according to rules and thus restrict their liberty, those who have submitted to these restrictions when required have a right to a similar submission from those who have benefitted from their submission" (1955, p. 185). This fair play-view represents a distinctive and compelling view on political obligation, i.e. the duty of citizens to accept the authority of the state and its commands (Klosko 1992, McDermott 2004).

While the fair play view can be extended to cover many cases of political obligation, it is particularly apt in the case of law and punishment. The system of law and order can easily be said to be a common good (Bird 2006, p. 162). In promoting security, safeguarding rights, and establishing predictable terms of social existence, it is to the benefit of all – *including* those who desire to break the laws. For even an offender is better off under a system of law and order. As Dennett has put it in another context, "even the rational psychopath will have an internal justification for supporting laws that punish psychopaths, since they protect him from other psychopaths" (Dennett 2003, p. 298). So when a person breaks the laws, she free rides: she enjoys the *benefits* of the system of law without accepting the fair share of the burdens needed for that system to prevail. She casts off the burdens of compliance with the laws and is thereby in a position superior to the law-abiding. This is unfair. The key claim of FPR is that punishment should aim to restore the fairness between offenders and the law-abiding and that it is justified to the extent it serves this end.

In being rooted in a plausible account of fairness and political obligation, FPR offers a compelling potential solution to the problem of punishment. This is not least so since it can easily accommodate the Rawlsian claim that principles of justice must respect the "separateness" of persons (1999, p. 163). A perennial weakness of consequentialist theories is that they seem unable to account for that individual have rights. Their focus on beneficial overall consequences can blind them to other factors that seem relevant to the permissibility of punishment, such as the idea that the offender has a right not to be treated (merely) as a means to social utility (Murphy 1973). FPR is readily compatible with the idea of people having rights, including the offender. Indeed, a plausible contractualist case can be made that fairness-restoring punishment is something which even the offender can in principle accept (cf. Scanlon 1998; Brettschneider 2007). This does not involve the *actual* attitudes of the person suffering punishment, of course. When we are evaluating a punishment, we instead ask whether it is called for by a rule that the offender, like everyone else, would have wanted in place if he or she had a prior say in the matter. In other words, we seek not what people actually want *ex post*, but what they would have been reasonable to want *ex ante* (Duus-Otterström 2010). Murphy's pioneering defense of FPR proceeds along these contractualist lines:

> On this theory, a man may be said to rationally will X if, and only if, X is called for by a rule that the man would necessarily have adopted in the original position of choice – i.e., in a position of coming together with others to pick rules for the regulation of their mutual affairs. [...] Thus I can be said to will my own punishment if, in an antecedent position of choice, I and my fellows would have chosen institutions of punishment as the most rational means of dealing with those who might break the other generally beneficial social rules that had been adopted. (1973, p. 230)

According to FPR, everyone could be said to endorse punishment as a way of promoting the fairness of the social venture. Punishment thus passes the contractualist criterion of following from principles which no one could reject *ex ante*.

3. Project description

FPR represents a distinctively retributivist,

backward-looking theory that is located snuggly within a wider politico-theoretical literature, and which, while it retains the characteristic focus on desert, explains in plausible terms why we are justified in acting on desert. FPR holds out a lot of promise as a solution to the problem of punishment.

Yet, for all its promise, FPR is insufficiently understood. More work is needed to flesh out and assess the theory. In particular, the relation between fair play obligations and punishment needs tight scrutiny. The theory must also meet a set of challenges laid down by its critics. Since I believe that defending the theory against its main objections is equivalent to understanding its positive claims, I here choose to focus on the objections that the theory needs to address:

OBJECTION 1: IRRELEVANCE

FPR holds that unfairness explains criminal desert: it is the fact that an offender takes unfair advantage of others that grounds that he or she is deserving of a punishment. The unfairness explains both why it is *permissible* to punish an offender, and why we normally *ought* to do so. Critics have held that FPR misrepresents the nature of criminal desert, arguing that while FPR can account for criminal desert when it comes to *mala prohibita* crimes, it cannot account for *mala in se* crimes (Duff 2001). The worrying thing about particularly heinous *mala in se* crimes seems to be that they are very harmful, not that they are unfair. The wrongness of murder, for example, seems to have more to do with the harm it causes the murdered and his or her loved ones than with it being unfair to the law-abiding (Dolinko 1991). This charge is thus mainly conceptual: the argument is that FPR cannot properly account for criminal desert. This is significant since FPR, as a species of retributivism, needs to get desert right, in particular because it would otherwise not get the ranking right between crimes and punishments, which is key for proportionality (von Hirsch et. al. 2009).

OBJECTION 2: FALSE EQUIVALENCE

Proponents of FPR have responded to the irrelevance objection by saying that while crimes can differ in terms of other moral characteristics, they all *also* include casting off the general burden of self restraint and hence involve unfairness (Boonin 2008, p. 124-126). The worrying thing about this so called *general compliance response*, however, is that it seems unable to *differentiate* between crimes. The general compliance response holds that all crime equally involve the same (general) unfairness of free riding. Yet this seems unable to capture some of our considered intuitions – for example, that murder is deserving of a much harsher punishment than tax evasion.

A related worry is that FPR, in saying that punishment is right because it removes the advantage the criminal has taken over the law-abiding, will say that the wrongness of a crime depends on whether the offender takes liberties that also tempt the law-abiding (Dolinko 1991). FPR says that the wrongness of crime stems from the offender's casting off the burdens of self-restraint that are shouldered by the law-abiding. But most people simply are not tempted to commit severe crimes. Hence, it is no burden for them to refrain from committing them. The conclusion seems to be that FPR will treat crimes as worse the more people are tempted to commit them. This has absurd implications for the way we rank crimes. A crime that tempts many (tax evasion) will be deemed worse than a crime that tempts very few (like murder). The upshot seems to be that FPR can at best only handle a subset of crimes, at worst is irredeemably flawed.

OBJECTION 3: THE UNEQUAL PLAYING FIELD RESPONSE

FPR locates the nature of our obligation to obey the laws in fairness: it is because people benefit from a law that they have a fair play duty to respect it. This stance, however, invites the objection that FPR seems to presuppose

just societies, i.e. societies where there is a level playing field in terms of access to advantage and where rules are to the benefit of all. Since actual societies arguably fail to meet this requirement, one can question whether FPR has any purchase as a solution to the problem of punishment. Some retributivists have argued that retributivism should not be wedded to contractualist reasoning precisely because actual societies are not fair (Anderson 1997).

There are two versions of this objection. First, it might be argued that some individuals do not benefit, on aggregate, from the system of law. For these individuals it seems that the contractualist test at the heart of the fair play account cannot be met (cf. Hoskins 2010). Second, it might also be argued that while everyone does benefit from the system of law, they benefit to vastly different degrees depending on their position in society. If the "cooperative venture" benefits some more than others – people with a certain ethnic or socioeconomic background, for example – then one might wonder whether the disadvantaged have a lesser obligation to obey the law.

It might be responded that FPR is an ideal theory and that we should not assume that actual societies have a right to punish (Murphy 1973; Dagger 2008). Yet if FPR is to handle real-world jurisdictions, the unequal playing field objection is clearly worrying. It serves to remind us that punishment should not be discussed without paying attention to the background conditions of the society which it is tasked to protect. As Honderich (2006) has noted this is a lacuna of the literature of punishment in general (but see Heffernan & Kleinig 2000). However, the lacuna is particularly troubling in the case of FPR in that this theory places fundamental importance in the fairness of the society.

Outline of the project

These are not the only objections to FPR, but they seem to me the most troubling ones. I have no settled view as to whether they can be overcome, and, if they cannot, whether this should make us on balance reject FPR. Recall that the project aims to answer two broad questions:

1. To construct, based on previous accounts of FPR and relevant branches of political theory, the best (in the sense most plausible) statement of fair play retributivism.
2. To assess the strength of that theory as an answer to the problem of punishment by way of analyzing a set of objections to it.

The objections point out the need to specify the most plausible version of FPR, as mentioned in (1). More specifically, this task can be broken down into a set of sub-questions:

1.1 What is the most plausible account of benefits and burdens with regard to the system of law?

1.2 What are the conditions that make punishment permissible?

1.3 What are the conditions that make punishment desirable or required?

The answer to 1.1 will in all likelihood turn out to be that the *benefits* of the system of law are primarily associated with providing individuals a sphere of autonomy and non-interference that allows them to pursue their conception of the good in relative security, while the *burdens* stem from having to refrain from doing things one would otherwise like to do. This opens up for the charge that the benefits of the system of law is "forced" – people by and large cannot exit the system(s) and hence cannot be said to freely accept its benefits (Simmons 1996). However, I follow Klosko in arguing that if benefits are open – they amount to public goods that benefit all regardless of whether one wants them – and exit is impossible, then there is a duty to reciprocate compliance with compliance (Klosko 1992). This response admittedly presupposes the plausibility of the fair play account of political obligations, and indeed the project will assume that this account is by and large correct.

Next, the project needs to analyze the

particular idea of a criminal's unfair advantage as paving the way for punishment. As is evident from 1.2 and 1.3, it is desirable to split this question into one involving the *permissibility* of punishment and one involving the *positive reason* to punish. Some construe FPR as a species of negative retributivism (Mackie 1982) that says that while criminal desert establishes the permissibility of punishing a particular lawbreaker, other considerations establishes both the reason to punish and the particular mode of punishment (Stichter 2010). It is thus not obvious that the notion of unfair advantage does work across the board, on all major questions of punishment (cf. Hart 2008). This idea is potentially very significant as it could overcome the irrelevance and false equivalence objections. The state might be justified in punishing murderers harder than people who cheat on their taxes not because they are unfair to a different extent, but because of the greater interest people have in having would-be murderers deterred. Thus, the project will investigate whether a partial reliance on social protection and overall welfare might serve to make FPR more attractive. This of course makes FPR belong to the set of "hybrid" theories.

Once the best statement of FPR is in place, the second overarching aim (2) is to assess the strength of that theory as an overall answer to the problem of punishment. This aim will require careful analysis of the attractions and flaws of that theory when compared with rival theories. The rival theories that seem most relevant here is a deterrence-based theory, intrinsic-good retributivism, and a communicative model of the kind currently defended by Duff (2001). The method used to assess the theories is *wide reflective equilibrium*-style normative analysis (Daniels 1979). This method, which is heavily favored by contemporary theorists, consists in seeking "fit" between one's considered ethical judgments in particular cases and general principles or theories. This is predominantly achieved by way of testing principles or theories against

examples (some real-world, other hypothetical) in order to see how they fare relative each other (McDermott 2010). The analytical work tends to attach priority to firmly held particular judgments: if a theory is shown to imply a highly counterintuitive action, this is usually taken as a reason to abandon or modify the theory. Yet the possibility also remains that the most attractive way of achieving fit is to bite the bullet and accept the implication, as per standard Rawlsian reflective equilibrium methodology (Rawls 1999: 40-46).

4. Significance

Punishment is arguably the most intrusive of domestic exercises of state power. The problem of punishment therefore deserves close and serious attention. Not least for the members of our communities that suffer punishment it is important to think about how their treatment can be justified to them, and how we need to modify the practice of punishment to make it cohere with the nature of that justification. For these reasons, the non-academic relevance of the issues pursued in this project is significant.

In terms of the significance for the academic field it is safe to say that the project's results will interest scholars working on punishment. While the fair play account continues to generate supporters and detractors, it is also insufficiently understood. Working out the issues (conceptual as well as normative) of the theory in a systematic fashion would interest not only those involved specifically in the debate over FPR, but also the wider community working on punishment in general. In particular, there are as of yet no book-length treatments of FPR. The project also aims to contribute to the literature on political obligations and the state responsibility to handle non-compliance. It is one of the merits of FPR that it so coherently analyses the practice of punishment as a *political* practice. Political theorists in general will benefit from the close attention the project pays to

fair play obligations and their connection to punishment.

5. Preliminary results

As stated above, in my Ph.D. thesis I defended a version of retributivism on contractualist grounds. While I am now convinced I got some things wrong, I am still convinced that a plausible theory of punishment will have significant retributive elements and must pass an *ex ante* contractualist test, i.e. be the object of reasonable consent even among the very people who suffer punishment. More specifically, I hold that a plausible theory of punishment must (a) explain why punishment is permissible on backward-looking grounds, (b) entail a reasonable proportionality between the seriousness of crimes and the harshness of punishment, and (c) pass a contractualist test. Whether FPR can deliver a coherent and compelling answer that ticks all three dimensions is an open question whose answer naturally must await the actual analysis. But it is fair to say that I am positively predisposed to FPR. I regard it as the most promising way of achieving these desiderata, with the focus on fairness yielding what seems like attractive answers on (a) and (c) in particular.

However, it is clear that the three objections mentioned above raise serious doubts about FPR. This holds in particular for the irrelevance and false equivalence objections.[1] What FPR gets right is no doubt that crime violates the social contract. Yet the general compliance argument is not a convincing answer to these worries, and the coherence of a mixed view such as the one recently espoused by Dagger (2008) is in doubt. I currently believe that the best way forward is to rethink the very role of fairness of the theory.

The thing that critics and advocates alike

have overlooked is that people might fare better than others *simply by making others fare worse.* Fairness is an essentially comparative notion. In Temkin's classic phrasing, commitment to fairness (or equality, in his wording) is to hold that "it is bad for some to be worse off than others through no fault or choice of their own" (Temkin 2003: 767). A lawbreaker no doubt often leaves her victims worse off through no fault of their own. It also stands to reason that the intuitively worst crimes – acts of physical or sexual violence, for example – leave the victims the worst off. One could thus say that while the violent criminal has not necessarily benefitted in an absolute sense from her crime, she most definitely stands in a relation of *unfair advantage* to her victims who are now faring much worse. The upshot seems to be that fairness can provide reasons to punish that are similar to compensatory or corrective justice. Punishment is simply the most natural way to restore basic fairness between persons, and the reason why the worst crimes merit the most severe penalties is that their victims are made the worst off.

This account is sketchy and raises several questions. First, we might wonder if it strays too far from the notion of free riding as an instance of unfair play. Second, like all comparative accounts it is susceptible to leveling down and that might seem a problem (see e.g. Knight 2009). I believe, however, that punishment is the best example of a situation where we *can* accept leveling down as appropriate. Third, it seems to offer counterintuitive results when it comes to victimless crimes such as tax evasion and cases of deadly violence (in the latter case we might feel that the victim simply ceases to exist and cannot be said to be "worse off" than others). Still, I believe that a distinctive and attractive account on FPR potentially could be constructed by mining the impressive general literature on egalitarianism and fairness. If not, and if no other account can offer compelling answers to the irrelevance and false equivalence questions, the project will simply argue against FPR

---

1   The unequal playing field objection does raise significant worries about FPR too, but as it concerns the viability of the theory "only" under non-ideal circumstances it is less fundamental

– a result as instructive and worthwhile as its opposite. Only time will tell which way the argument goes.

References

Anderson, J. L., 1997. "Reciprocity as a Justification for Retributivism", *Criminal Justice Ethics* 16:1, pp. 13-26.

Bird, C., 2006. *An Introduction to Political Philosophy*. Cambridge: CUP.

Bentham, J., 1988. *The Principles of Morals and Legislation*, New York: Prometheus Books.

Berman, M., 2008. "Punishment and Justification", *Ethics* 118: 258-290.

Boonin, D., 2008. *The Problem of Punishment*. Cambridge: CUP.

Brettschneider, C., 2007. "The Rights of the Guilty: Punishment and Political Legitimacy", *Political Theory* 35: 175-99.

Brå (Brottsförebyggande rådet), 2010. *Kriminalstatisik 2009*. Rapport 2010: 15.

Dagger, R., 1993. "Playing Fair With Punishment", *Ethics* 103: 473-488.

Dagger, R., 2008. "Punishment as Fair Play", *Res Publica* 103: 473-488.

Daniels, N., 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics", *The Journal of Philosophy* 76: 256-282.

Davis, M., 1993. "Criminal Desert and Unfair Advantage: What's the Connection?", *Law and Philosophy* 12: 133-156.

Dennett, D., 2003. *Freedom Evolves*. New York: Penguin Books.

Dolinko, D., 1991. "Some Thoughts About Retributivism", *Ethics* 101: 537-59.

Duff, A., 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press.

Duff, A., 2008a. "The Incompleteness of 'Punishment as Fair Play'", *Res Publica* 14: 277-281.

Duff, A., 2008b. "Legal Punishment", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall2008/entries/legal-punishment/

Duus-Otterström, G., 2007. *Punishment and Personal Responsibility*. University of Gothenburg: Dept of Political Science.

Duus-Otterström, G., 2010. "Fallibility and Retribution", *Law and Philosophy* 29: 337-369.

Hanna, N., 2009. "Liberalism and the General Justifiability of Punishment", *Philosophical Studies* 145: 325-349.

Hart, H.L.A., 1995. "Are There Any Natural Rights?", *The Philosophival Review* 64: 175-91.

Hart, H.L.A., 2008. *Punishment and Responsibility* (2nd ed.). Oxford: Clarendon Press.

Heffernan, W. C. & J. Kleinig (eds), 2000. *From Social Justice to Criminal Justice*. New York: Oxford University Press.

Honderich, T., 2006. *Punishment. The Supposed Justifications Revisited*. London: Pluto Press.

Hoskins, Z., 2010. "Fair Play, Political Obligation, and Punishment", *Criminal Law and Philosophy* 5: 53-71.

Klosko, G., 1992. *The Principle of Fairness and Political Obligation*. Lanham: Rowman and Littlefield.

Knight, C., 2009. *Luck Egalitarianism*. Edinburgh: EUP.

Kymlicka, W., 2002. *Contemporary Political Philosophy* (2nd ed.). Oxford: OUP.

Mackie, J.L., 1982. "Morality and the Retributive Emotions", *Criminal Justice Ethics*, pp. 3-10.

McDermott, D., 2004: "Fair-Play Obligations", *Political Studies* 52: 216-232.

McDermott, D., 2010. "Analytical Political Philosophy", in Leopold & Stears (eds.), *Political Theory: Methods and Approaches*. Oxford: OUP.

Moore, M., 1998. *Placing Blame: A General Theory of the Criminal Law*. Oxford: Clarendon Press.

Morris, H., 1968. "Persons and Punishment", *The Monist* 52: 475-501.

Murphy, J., 1973. "Marxism and Retribution", *Philosophy & Public Affairs* 2: 217-243.

Rawls, J., 1999. *A Theory of Justice* (revised ed.). Oxford: Oxford University Press.

Rawls, J., 2001. "Two Concepts of Rules", in *John Rawls: Collected Papers*, edited by Samuel Freeman. Cambridge, MA: Harvard University Press.

Scanlon, T.M., 1998. *What We Owe to Each Other*. Cambridge: HUP.

Shafer-Landau, Russ, 1996. "The Failure of Retributivism", *Philosophical Studies* 82: 289-316.

Simmons, A.J., 1996. "The Principle of Fair Play", in Robert (ed.) *Readings in Social, Political Philosophy*. Oxford: OUP.

Stichter, Matt, 2010. "Rescuing Fair Play as a Justification for Punishment", *Res Publica* 16: 73-81.

Temkin, L., 2003. "Egalitarianism Defended", *Ethics* 113:764-782.

von Hirsch, Andrew, Ashworth, Andrew & Roberts, Julian (eds), 2009. *Principled Sentencing* (3rd ed.). Oxford: Hart Publishing.