

Nicolas Nayfeld. *Moral Pluralism and the Complexity of Punishment: The Penal Philosophy of H.L.A. Hart.* Routledge 2023. 218 pp. \$160.00 USD (Hardcover 9781032271224).

The debate over the ethics of criminal punishment is in full gear these days, and it is remarkable what a wide range of options there are for different theories, or even for those who insist that criminal punishment is unjustified. Back in the 1960s, one of the most influential theories of punishment was by Hart, and it continues to have its supporters today. Nayfeld's monograph provides us with a systematic survey, interpretation, and general defence of Hart's penal philosophy. The author is well-informed on Hart's work and willing to criticize or supplement it where it seems wrong. As the author demonstrates, Hart's theory remains an important contribution to the debate, though Nayfeld underestimates the problematic nature of Hart's account.

One of the author's main contentions is that Hart's theory should not be seen as a 'mixed theory' of punishment but rather as a 'pluralist' theory. Mixed theory is a broad approach for justifying punishment that became highly influential in the 1950s and 1960s. It attempted to reconcile the two competing theories of punishment, utilitarianism, and retributivism, by distinguishing two distinct 'levels' or 'questions' regarding punishment. Utilitarianism would provide the rationale for punishment, while retribution would provide the constraints on that theory, ensuring that only the guilty are punished and only proportionately (sometimes called 'negative retributivism').

Nayfeld insists that Hart was not a 'mixed' theorist since Hart rejected retributivism. This claim is, however, not convincing. For one thing, not all mixed theorists straightforwardly accepted the retributive theory; some used a 'linguistic' strategy, claiming that the very logic of punishment implies that only the guilty be punished and only proportionately. Hart famously criticized this argument as the 'definitional stop.' However, Hart adopts the same strategy, claiming that punishing for past wrongs is a 'semantic fact.' Nayfeld surprisingly gives Hart a pass on this inconsistency (6).

Equally important, Nayfeld misses what is distinctive about a mixed versus a pluralist theory. A mixed theory is a version of pluralism that tries to avoid conflict by segregating the different theories into distinct aspects or 'questions.' The linguistic idea is one such strategy: utilitarianism provides the purpose of punishment, but language provides the deontological constraints. Hart also adopts the famous distinction between the 'general justifying aim' of punishment (which is utilitarian) and the 'distribution' of punishment --who gets what punishment. The latter is not based on utilitarianism, though Hart is less than clear about its basis. Hart tells us it is a principle of 'retribution in Distribution,' calling into question Nayfeld's insistence that Hart rejects retributivism (Hart rejects only positive retributivism).

So, Hart's theory is clearly a mixed theory rather than a true pluralist theory. A pluralist theory would hold that there is no substantial distinction between different 'questions' and that the justification of punishment cannot be divided into the 'general justifying aim' and the 'distribution,' a highly artificial separation in which, oddly, two different moral theories are at work in different 'aspects' of the punitive system. The very awkwardness of Hart's terminology suggests just how strange this distinction is. What is a 'general justifying aim'? As distinct from a 'particular' justifying aim? Is the justification limited to the aim or purpose? Moreover, who would ever describe the



sentences imposed as the ‘distribution’ of punishments? A pluralist theory would never hold that the ‘general justifying aim’ of punishment must be solely utilitarian; for a pluralist, the aims and goals of punishment are multiple and varied and need to be weighed against each other.

Nayfeld argues that it is a virtue of Hart’s analytic method that he can make important conceptual distinctions such as this, and that other theorists have simply ‘conflated’ the two questions. It seems more likely that these distinctions are not only artificial, but indeed invented simply as a strategy to reconcile the competing theories of punishment. Indeed, Nayfeld recognizes in various places that the distinction between ‘general justifying aim’ and ‘distribution’ breaks down; he admits for example that the two are ‘interdependent’ (20) but does not seem to think that this admission is in any way problematic. But most philosophers have rejected Hart’s distinction as untenable and implausible. In what sense is there a ‘logical priority’ of the ‘general justifying aim’ over the ‘distribution’ (67)? Surely, the ‘distribution’ of punishment is crucial to justifying a penal system. Nayfeld uses the analogy of slavery, claiming that slavery is intrinsically morally wrong, regardless of its ‘distribution’ (i.e., who is enslaved and how they are treated) (67). However, the analogy is unconvincing. Slavery is wrong because it treats people as a means (the Means Principle). But one cannot reach this conclusion by simply analyzing the (utilitarian) general justifying aim of slavery, as utilitarians do not recognize the Means Principle.

Just what the purpose of the book is, is somewhat unclear. Is it meant to be an exposition, a reinterpretation, or a general defence of Hart’s theory? The author seems to think that Hart’s theory remains a viable moral justification for punishment, albeit with some additions and revisions. Take, for example, the classic objection to the utilitarian theory: that it violates the Means Principle. Nayfeld quickly dismisses this objection by claiming that punishment satisfies the Means Principle so long as individuals have the free choice whether to commit a crime or not (93). Nevertheless, the ‘general justifying aim’ is supposed to be utilitarian, and it does not recognize the Means Principle. In any case, the argument is far too quick; indeed, it would permit wildly disproportionate punishments (death penalty for jaywalking).

Nayfeld identifies as ‘one of the most important principles of Hart’s penal philosophy’ that ‘the distribution of punishment cannot be inferred from its general justifying aim and vice versa’ (169). This is a good statement of the heart of Hart’s theory and why the theory is, to most philosophers, untenable. It is a striking feature of the debate that followers of Hart take this distinction as too obvious to need defense, while most moral philosophers (including myself) think that the distinction is wholly artificial and distorts the debate. The idea that the justifying purpose of punishment should have no relevance to the application of punishment to individuals or that the punishment meted out to individuals should have nothing to do with the justifying purpose of punishment, seems not wrong but bizarre. If the purpose of punishment is to prevent crime, how could it be the case that the decision of whom to punish and how much should have nothing to do with crime prevention? If all Hart means is that the ‘distribution’ of punishment should involve multiple competing values, that is all well and good. But then those competing values should also be reflected in the ‘general justifying aim.’ The idea of a ‘logical’ distinction between the two does not hold up.

This book provides a valuable exposition of Hart’s penal philosophy for anyone interested in

Hart on the philosophy of punishment. It should, however, be read with a critical eye given that the author seems rather too accepting of some of the problematic aspects of Hart's theory.

W.R. Kaufman, University of Massachusetts