Meeting 8: Chapter 4, 1st Half (Nozick on Tracking Value and Retributive Punishment)

I. Does Evolutionary Theory Undercut Tracking?

According to Nozick, S’s doing of act A tracks rightness if and only if: (1) A is right; (2) S does A; (3) if A weren’t permitted, S wouldn’t do A; and (4) if A were required, S would do A.

A threat from evolutionary theory:

Evolutionary theory offers us an “invisible hand” explanation of ethical behavior in terms of the degree to which it increases inclusive fitness and thus, when genetically based and heritable, over generations spreads more widely: “Ethically behaving individuals will leave more great-grandchildren or (given kin selection) great-grandnephews and nieces similarly disposed” (p. 344).

This explanation appears to explain our ethical behavior without assuming the truth of any facts about rightness.

So, it seems, our ethical behavior does not track rightness.

We can argue in a similar way that our ethical beliefs (e.g. our beliefs about rightness) do not track the truth, since there is an evolutionary explanation of them that does not appeal to the truth of what is believed.

This appears to yield a very strong conclusion:

“This type of [evolutionary] explanation does more than threaten the link wherein our beliefs track [ethical] facts; if successfully carried through, it also would seem to show that it is unreasonable to believe there are any such [objective] [ethical] facts at all. . . . Given the absence of any demonstration or strong argument that there are such facts, despite the repeated attempts by philosophers, if also all of our ethical behavior and beliefs can be explained without in any way assuming their existence, then it would seem unreasonable to continue to think there are these facts” (p. 342)

Nozick’s initial response: “It is compatible with an evolutionary explanation of a capacity that it is a capacity to detect or track some truths” (p. 345).

His example: an explanation of the adaptive value of a capacity to form rudimentary arithmetical beliefs (such as a belief that 2 + 2 = 4) must appeal, in part, to the truth of those beliefs.

So although the spread of this capacity is explained “blindly,” the capacity itself counts as a capacity to recognize and respond to arithmetical truths and hence is not blind at the individual level.

Moreover, “the capacities of recognition selected for will reveal splendors beyond the then useful. The abilities . . . to recognize, prove, or understand a proof that there is no greatest prime number . . . made no special contribution to survival. But the mathematical realm, once entered for narrow utilizable purposes merely, exhibits its own further structures and intricacies” (p. 346).

Objection: but the ethical case is different, because here we can explain the adaptive value of ethical beliefs and behavior by appealing to the non-normative facts in virtue of which the ethical facts hold.

Nozick’s reply (which is now known as a third-factor response to evolutionary debunking arguments):

“[E]thical behavior will serve inclusive fitness through serving or not harming others, through helping one’s children and relatives, through acts that aid them in escaping predators, and so forth; that this behavior is helpful and not harmful is not unconnected to why (one most theorist’s [sic] views) it is ethical. [E]thical behavior will increase inclusive fitness through the very aspects that make it ethical, not as a side effect through features that only accidentally are connected with ethicality” (ibid.).
II. Explaining the Deservedness of Retributive Punishment
Although Nozick does not think that the topic of punishment is central to the free will problem, he does believe that punishment in retribution for a wrongdoing can, in certain cases, be appropriate or deserved.

So in Part III, he turns to the question, “How is it possible that retributive punishment is sometimes deserved?”

What does this Part have to offer those of you who think such punishment is never deserved?

First, this is a case where an explanation of how X is possible seems also to constitute a justification of X, so anti-retributivists can see Nozick’s explanation as offering them an argument they must rebut.

Second, anti-retributivists can try to turn Nozick’s explanation of deserved retributive punishment into a reductio of the notion, by arguing as follows: if there is deserved retributive punishment, this is the best explanation of how it is possible; but that explanation is not plausible; so there is no such thing.

Can we just hold that the appropriateness of retributive punishment is fundamental/unexplained?

No, insists Nozick: the retributivist position is “not . . . smoothly shaped,” with “surprising contours” and “irregular dips and angles.” It is implausible “that fundamental facts having no further explanation would take that shape” (p. 366).

III. The Explanandum
Suppose a person, S, inflicts a wrong of magnitude H on others.

Should this be the level of wrong intended or the level of wrong done? Nozick thinks H either should be an average of the two or should be the level of wrong done, unless the level of wrong intended is greater, in which case H is the average of the two.

Let r (at first) be S’s degree of responsibility for the act.

r varies from 1 (full responsibility) to 0 (no responsibility).

Let ε be the degree of harm inflicted on S in the process of making compensation to the victims of S’s act.

Nozick’s proposal (p. 364): S deserves a punishment of degree \( r \times H - \epsilon \), if this quantity is greater than 0.

Is this a primitive view, expressive only of a thirst for revenge?

No, says Nozick, because retribution and revenge are different in the following ways (pp. 366-68):

1. Retribution is done for a wrong, while revenge can be done for a mere injury, harm, or slight.

2. Retribution sets an internal limit to the amount of punishment inflicted, according to the seriousness of the wrong, while revenge need set no internal limit to what is inflicted.

3. Revenge requires a personal tie to the victim, while retribution does not.

4. Revenge involves an emotional tone, viz. pleasure in the suffering of another, while retribution either involves no emotional tone or involves another one, viz. pleasure at justice being done.

5. Retribution has a generality: it commits one to enacting similar punishments in other similar circumstances, whereas revenge has no such generality.

Later Nozick reinterprets r to be “the degree to which the person flouts correct values, except for that component of his degree of flouting included in H” (p. 388).

So interpreted, r can have a non-zero value even if the wrongdoer’s action is causally determined.

Note, finally, that just because a certain type of punishment is deserved, it doesn’t necessarily follow that we should have an institution that attempts to enforce that type of punishment.
IV. Retributive Punishment as a Message

In “the fullest and most satisfactory case” of a retributive punishment for S’s act A, we have (p. 369):

1. Someone believes that (S’s act A is wrong to a certain degree)
2. and visits a penalty upon S
3. which is determined by $r \times H - c$,
4. intending that (the penalty be done because of wrong act A
5. and in virtue of the wrongness of A),
6. intending that S know that (the penalty was visited upon him because he did A
7. and in virtue of the wrongness of A,
8. by someone who intended to have the penalty fit and be done because of the wrongness of A
9. and who intended that S would recognize the penalty was visited upon him so that 1-9 are satisfied).

Why we need 5: “If S wrongfully shoots another in a canyon and the sound of the shot causes an avalanche that maims or kills S, then this happens to S because of his wrong act but not because of the wrongness of the act” (ibid.).

Why we need an intention in 4: “If, on the cliff above, a witness sees the wrongful act and scrambles off to get forces of the law, thereby kicking loose some stones that cause an avalanche, the ensuing crushing of the killer still does not occur in retribution for his act” (ibid.).

This complicated structure, “wherein something intentionally is produced in another with the intention that he realize why it was produced and that he realize he was intended to realize all this” (p. 370), fits an influential account of meaning due to H. P. Grice.

So, Nozick concludes, retributive punishment sends a message to the person being punished, namely: “This is (at least) how wrong what you did was.”

V. Teleological Retributivism

One form of retributivism focuses on the recipient of this message.

According to teleological retributivists, the goal of retributive punishment is to get the wrongdoer to realize the wrongness of his act, by bringing home to him the nature of what he has done.

Two advantages of teleological retributivism:

- It can explain our reluctance to punish a “person who is mentally defective so as to be incapable of learning or realizing that his act was wrong” (p. 372).
- It can explain our reluctance to punish “someone who already realizes his act was wrong and is repentant, attempting to make amends, and so forth” (ibid.).

Nozick’s primary objections to teleological retributivism (as the sole theory of deserved retributive punishment):

- Teleological retributivists should welcome alternative, less harmful routes to their goal of moral learning. But “[e]ven were it possible to produce in him the realization that he acted monstrously and evilly, through his seeing films, reading novels, and hearing explanations of the causes of his behavior and the tales of his victims, but with tranquillizers administered to prevent his suffering at the realization of the enormity of what he had done, would we really want merely this to have been done to Adolf Hitler, had he been capture alive?” (p. 373).
- It appears that teleological retributivists should be reluctant to punish people who are too “hardened and resistant to recognize[c] that they acted wrongly” (p. 374).
VI. Non-Teleological Retributivism

Nozick prefers a form of retributivism that focuses on the transmission of the message itself.

The basic idea: “a wrongdoer has become disconnected from correct values, and the purpose of punishment is to (re)connect him,” via that very act of retributive punishment (p. 374).

Why do we do this? “Correct values are themselves without causal power, and the wrongdoer chooses not to give them effect in his life. So others must give them some effect in his life, in a secondary way” (p. 375)

Nozick’s non-teleological retributivism, in a nutshell: “Retributive punishment is to effect two things: (a) connect the wrongdoer to value qua value (b) so that value qua value has a significant effect in his life, as significant as his own flouting of correct values” (pp. 376-77).

Condition (a) is why the Gricean intentions are needed: “The complicated (Gricean) intentions enable us to act as a vehicle whereby correct values, qua correct values [i.e. in their nature as correct values], act upon the wrongdoer” (p. 376).

Condition (b) is why a punishment is needed: “The punishment is central—that is the way the correct values which he has flouted have a significant effect on his life” (ibid.).

Punishment that gives correct value, qua correct value, a significant effect on the wrongdoer’s life is good in itself, claims Nozick.

But, he allows, the sort of moral improvement that teleological retributivists seek is also a good thing (though it is an extra bonus that we sometimes get, not the primary rationale for punishment).

Does it follow from teleological retributivism that capital punishment is sometimes justified?

Why this might follow: for the most serious floutings of the most important values (e.g. the intentional murdering of another), capital punishment would appear to be a response of equal magnitude.

(This distinguishes Nozick’s view from teleological retributivism, which has a hard time justifying capital punishment.)

Why this might not follow: in imposing capital punishment on the wrongdoer, we might be incorrectly responding to his value. “Some acts, even if they effect his connection with correct values, may snap your connection with his value, so as to constitute your flouting his value” (p. 378).

If the rationale for retributive punishment is to effect a connection with value qua value, why do we only punish wrongful acts? Why not punish people who haven’t done wrong but would if given the chance?

An external explanation of the act requirement for punishment: people have a right to be left alone, which they lose when they commit wrongful acts, and only then may a punishment be imposed.

Why Nozick is dissatisfied with this explanation: then non-teleological retributivism loses an important advantage over deterrence and reformative theories of punishment, since they can appeal to the same external explanation to fend off standard objections to their view.

An internal explanation of the act requirement: the person who flouts correct values is not just unlinked with those values, but rather anti-linked with them. And it is through actions that correct values are flouted. “Since it is flouting that is to be punish, actions are required” (p. 384).

But isn’t a person who is disposed to do evil qua evil, but hasn’t acted on this disposition, anti-linked with value?

Nozick’s reply: “Anti-linked perhaps, but not flouting” (p. 384). But isn’t this to take back the claim that flouting matters because it’s a way of being anti-linked with value?