Sanction and Obligation in Hart’s Theory of Law

DANNY PRIEL

Hart used his famous example of the “gunman situation” to demonstrate that legal obligations can exist even without sanctions. I will argue that Hart’s argument is unsuccessful.

Right at the start of *The Concept of Law* Hart says that “[t]he most prominent general feature of law at all times and places is that its existence means that certain kinds of conduct are no longer optional, but in some sense obligatory” (Hart 1994, 6).¹ But like law, the gunman situation also involves conduct which is in some sense is obligatory. Hart therefore had to show that even though both situations involve a person who is required to act or refrain from acting in a particular way—in both cases he is asked to comply with something I shall henceforth call a “demand”—these two situations are obligatory in different ways. Hart did that by drawing a distinction that has since become famous: whereas in the gunman situation the person subject to the demand may be obliged to give up his or her money, in the case of law one is under an obligation to act in a particular way.

Hart’s argument appeals to our intuitive judgment that the person subject to a robber’s threat is in a qualitatively different situation from the person subject to a legal demand. Whether this intuition is correct, and indeed the extent to which we should base philosophical arguments on intuitions at all, are both controversial matters,² but

---

¹ All further unidentified parenthetical references are to *The Concept of Law*.
² For challenging Hart’s intuition see Kramer (1999: 92-101). For arguments against the reliance on intuitions in philosophy see Hintikka (1999); for a jurisprudential echo see Leiter (2003: 43-51).
they are not questions I will address here. If we accept Hart’s methodological assumptions, it seems plausible to use this distinction to argue, as he does, that legal obligations exist not in virtue of the threat of sanction. As he puts it, “[i]n the case of the rules of the criminal law, it is logically possible and might be desirable that there should be such rules even though no punishment or other evil were threatened” (p. 34).³ The difference between the two can be put thus: in the case of the gunman, it is the threat and only the threat that explains why “conduct is in some sense obligatory”. In the case of law, however, there is something else, which exists independently of the threat that makes the action non-optional. It is true that in the case of law people may expect that failure to comply with the obligation will bring about some kind of censure or sanction. But this expectation is hardly necessary: According to Hart legal obligations continue to exist even when we know that they are violated and no sanction is likely to follow, as in the case of driving at a red traffic light in the middle of the night on a deserted street. Instead, what distinguishes the case of the threat and the case of the genuine obligation is that they involve utterly different attitudes on part of the person subject to the demand: in the case of the gunman what explains compliance with the demand is either action out of irrational fear or cold calculation of the relative costs and benefits of breaking or complying with the law. In contrast, obligations are perceived as reasons for action,

³ He adds immediately that this might not be a “legal rule”, but it will be a rule nonetheless (p. 34). However, later in the book he argues generally that “[t]o argue that international law is not binding [i.e., does not create obligations] because of its lack of organized sanctions is tacitly to accept the analysis of obligation contained in the theory that law is essentially a matter of orders backed by threats. … Yet … this identification distorts the role played in all legal thought and discourse of the ideas of obligation and duty” (pp. 217-18). And this, he adds, is true of municipal law as well. He gives additional examples in the endnotes (pp. 291-92).
perhaps even exclusionary reasons for action, which act as bars to usual cost-benefit reasoning.

This point serves to bolster Hart’s methodological point about the importance of examining law from the “internal point of view”, for he believed that this difference between threat and obligation can only be grasped by paying attention to the “internal aspect of rules” as seen from the perspective of those who take law to give them reasons for action. Only from this perspective can we see that the two situations that to an external observer might look identical are in fact quite different.

All we have said so far about legal obligations is, however, true of all obligations, and therefore is insufficient if we want to know what is unique to legal obligation. The short answer to this question is that an obligation can count as legal only if it is issued by a law, and a law can only exist within a legal system. So, based on Hart’s theory, for a legal obligation to exist the following are the minimal conditions must obtain:

1. Legal system $L$ exists.
2. There exists a valid (primary) rule in $L$ that imposes an obligation if a certain set of facts $F$ obtains.
3. $F$ obtains.4

What interests me are the conditions under which (1) obtains, i.e. the conditions under which a legal system exists. This, I believe, summarizes Hart’s view:5

---

4 I disregard here the question of “defeasibility” of legal rules. Hart argued (1949: 172-83) that the conditions in which a legal rule may be defeasible are distinct from the conditions that specify when it obtains. Later in life Hart rejected various aspects of this paper, but not, I think, the notion of defeasibility. In any case, all this is irrelevant for our concerns, so I leave it out here.

5 “There are … two minimum conditions necessary and sufficient for the existence of a legal system. On the one hand, those rules of behaviour which are valid according to the system’s ultimate criteria of validity must be generally obeyed, and, on the other hand, its rules of change and
(1.1) There is a certain (social) rule of recognition in a particular society.

(1.2) There are other secondary rules recognized by the rule of recognition.

(1.3) There are certain primary rules recognized by the rule of recognition.

(1.4) The rules of the systems are generally obeyed (although not necessarily accepted).

These conditions can be further analyzed thus:

(1.1.1) There exists a social rule specifying how to identify legal norms.

(1.1.2) The rule specified in (1.1.1) is accepted\(^6\) by legal officials.

(1.2.1) There exist power conferring norms specifying how the rules in (1.3) are to be changed and adjudicated.

(1.2.2) The norms specified in (1.2.1) are accepted by legal officials.

(1.3.1) There exist certain prescriptions identified by the rule in (1.1.1).

(1.3.2) The prescriptions specified in (1.3.1) are accepted by the legal officials.\(^7\)

Since Hart was trying to articulate the minimal conditions for the existence of a legal system he was clear that most members of society need not accept the primary rules; in fact, they do not even have to understand the state's rule of recognition: “surely the reality of the situation is that a great proportion of ordinary citizens—perhaps a majority—have no general conception of the legal structure or of its criteria of validity”, and that “the laws which are valid by the system’s tests of validity are obeyed by the bulk adjudication must be effectively accepted as common public standards of official behaviour by its officials. The first condition is the only one which private citizens need satisfy…. The second condition must also be satisfied by the officials of the system” (pp. 116-17).

\(^6\) In this context acceptance means the disposition to consider certain materials as the places for looking for the law, as well as (perhaps) the correct way to interpret them.

\(^7\) Here acceptance of a legal rule means that the rule is accepted as a sufficient reason for action, i.e. it is considered an appropriate answer to say that one should act in a certain way “because it is the law”.

---

\(6\) In this context acceptance means the disposition to consider certain materials as the places for looking for the law, as well as (perhaps) the correct way to interpret them.

\(7\) Here acceptance of a legal rule means that the rule is accepted as a sufficient reason for action, i.e. it is considered an appropriate answer to say that one should act in a certain way “because it is the law”.
of the population … surely is all the evidence we need in order to establish that a given legal system exists” (p. 114, to the same effect pp. 116-17).

Notice that (1.4) only demands that legal rules be obeyed, and (1.3.2) only requires acceptance of the primary rules by legal officials. This is consistent with two scenarios: one in which the primary rules are in fact accepted, and obeyed because accepted;\(^8\) the other, in which they are not accepted but obeyed because of fear of sanction. Because Hart was concerned with the minimal conditions for the existence of law, he was clear that all that was necessary for a legal system to exist is that the officials take the internal point of view (p. 203).

It follows that the existence of a legal obligation is different from the existence of a person feeling obligated. For a legal obligation to exist it may be that all those subject to law other than the officials think of their regime as an oppressive one, and therefore as essentially no different from the gunman situation writ large. A person may be under legal obligation even when he thinks he is in a gunman-like situation.

All this shows, first, why Hart’s contrast between genuine obligation and the gunman situation, despite all the work it was supposed to do in Hart’s theory, is misleading: the gunman situation focuses on the person to whom the demand is directed, whereas in the case of the legal system the focus is on the body making the demand. We have seen that in order for a person to be under a legal obligation all that has to obtain is that certain others accept the demand, even when the person to whom it is directed thinks of himself as being forced to do something by the threat of sanction. What distinguishes the gunman and the law must therefore be something other than the attitude of the

\(^8\) Incidentally Hart believed this to be the case in fact. He claimed that the external point of view fails to account for “the way in which the [legal] rules function as rules in the lives of those who normally are the majority of society” (p. 90, also p. 92).
person subject to the demand, and the only possibilities are the attitude of the person making the demand and the content of the demand. And since we can imagine a situation in which the content of the demand the law makes is similar to the one made by a robber, the only plausible possibility is the former, namely that what distinguishes law from the gunman situation is the attitude of the person making the demand.9

But if that is the case, it is a mistake to rule out the possibility that the gunman situation is one of law, for it depends on the attitude of the gunman: a robber with the attitude of Robin Hood might now be considered a legal system writ small. And if not, it is not because of a difference in the demand the robber makes, but because the robber situation lacks other features we usually find in a legal system (e.g., comprehensiveness, geographical boundaries, its mechanisms for change and adjudication).

But there is a deeper problem. (1.4) shows that Hart’s claim, that for legal obligation to exist all that is required is that officials take the internal point of view towards rules, is false. The problem is this: one of the features distinguishing an existing legal system from a hypothetical or defunct one is that the real one is generally obeyed (even if not generally accepted). But when it is generally obeyed but not generally accepted it has to resort to threats in order to maintain its efficacy. But from this it follows that in the “minimal” case of a legal system the threat of punishment is a necessary condition for the existence of a legal system, even if not in the straightforward way Hart rejected. The reason is not hard to see: even if we accept Hart’s point that an obligation may exist even

9 Could we say that the person subject to law would have seen the difference between law and the gunman’s demand had he taken the right point of view, and that he only fails to make the distinction because he is, say, Holmes’s “bad man”? Plainly not, for such an argument would be circular. The question that interests us is whether there is a difference between the gunman’s demand and the law’s demand, something we cannot show by assuming its existence.
when it is not obeyed, for a *legal* obligation to exist it has to be part of an active legal system. In a legal system in which only the officials accept the rules—a real possibility according to Hart—the legal system will remain active only because of the threat of punishment for disobedience. In other words, sanctions may perhaps not be required to show that a particular demand is an obligation, but they will be required to show that a demand is a *legal* obligation, because only in this way will the demand be part of an active legal system.

Hart can avoid this conclusion by saying that legal systems exist only when most people subject to them accept (and not merely obey) the rules. But this will undermine another part of Hart’s analysis: different people have attitudes with regard to what the law requires, but only some, the officials, have an authoritative say on the matter. This means that even when the laws are accepted by most people, we would need to know the officials are in order to have an authoritative judgment as to what counts as the law. But would we know that? They have to be either accepted as such or obeyed due to their use of force. When the entire community (or a substantial majority of it) accepts them as the officials, there is no problem with recognizing them; but when they are not accepted, then the only thing in virtue of which they will be considered as the officials of the legal system is their use of force and threat of punishment. It may be that a state in which the officials maintain their position by the threat of sanction is one found in the outer edges of what we call “law”. Indeed, some have doubted whether such regimes have what may be called a legal system, but Hart clearly did not belong to their ranks. As we have just seen, however, within such legal systems it is only the actual threat of sanctions that would allow the officials to maintain their position and thus keep the legal system alive.
If all this is true, what is the moral to be learned from it? One might be tempted to suggest that all we need to do is tinker with Hart’s account and revise it until we get a satisfactory model. For example, we may say that law exists either when it is accepted by most members of society, or when it is backed up by sanctions in order to cover those examples of law in a dictatorship. Even if such a patch were successful in keeping the definition standing, such an ad hoc approach seems arbitrary and it undermines the whole purpose of the Hartian enterprise of providing a unified answer to the question “what is law?”

Based on these considerations I believe the real conclusions to draw from this point—and others like it—are different. The history of such conceptual endeavours suggests that this project is not likely to succeed, and the argument above suggests why attempts at refining Hart’s account are likely to fail. Many contemporary legal philosophers, especially legal positivists, are trying to articulate the minimal conditions for the existence of law or a legal system. The assumption is that by finding those minimal conditions we will also find the features that all legal systems possess. The argument above challenges the claim that understanding actual, reasonably functioning and generally just legal system depends on giving an account of the minimal legal system, or to put it differently, that what distinguishes the just legal system from the unjust one is the contingent fact that the latter happens to have good (just, efficient, reasonable) laws.

When Lon Fuller presented his famous eight desiderata that “make law possible” (Fuller 1969: ch.2) legal theorists have replied that these desiderata only show what we are likely to find at an efficiently functioning legal system, whether good or evil (e.g., Kramer 1999: 67-71). I find these arguments unpersuasive (Priel: 2006: 116-17), but we
need not address this question now. What they miss is a methodological point that may be gleaned from Fuller’s ideas, and which, I think, is of my greater importance: the system of governance found in totalitarian regimes is likely to be a different beast from the one in non-totalitarian regimes, because the political values that underlie them are fundamentally different. A legal system that complies with Fuller’s rule-of-law desiderata are not the same as all other legal systems, only better at doing their job. A legal system that satisfies those conditions is not merely more efficient: it takes a particular view about moral responsibility, about agents’ ability to control their lives and the importance of letting them do that, it takes a position with regard to the relative position of citizens and their state, and it is committed to the view that it is important to give people adequate opportunity to plan their lives. This, I think, is the important point to take from Fuller’s account. Unfortunately, many legal theorists have interpreted this argument as one about the boundaries of law and non-law and whether evil regimes have law, issues that I believe only distract us from more important questions.

The right moral from Hart’s distinction between being obliged and being under obligation should be the same. The important question is not whether law is the gunman situation writ large or Don Corleone headed a legal system writ small. The question is what reasons underlie this distinction and explain why we make it. If there are none, then the distinction is arbitrary. Even if true, then, it represents nothing that legal theorists should spend much time arguing about. On the other hand, if there are such reasons, then they suggest that legal systems in dictatorial regimes are (or at least are believed to be) something different from those in non-dictatorial regimes not (just) in what they demand, but (also) in the way they make the demand. Understanding this issue would probably give us much greater insight into the traditional questions of
jurisprudence, but contrary to the view of many contemporary legal philosophers, trying to articulate those reasons and trying to understand why they are taken to be significant are a matter for political theory.¹⁰

University of Warwick
School of Law
Coventry CV4 7AL, UK

¹⁰ This conclusion may seem obvious, but is worth spelling out because it has been resisted for a long time by legal positivists (including most famously Hart himself in the Postscript to The Concept of Law), who have argued that their theory is morally neutral and descriptive (Marmor 2006; Gardner 2001: 202-03; Hart 1994: 239-40). For a different argument against this view see Priel 2007.
REFERENCES


