On Coercion and the (Functions of) Law

Julieta A. Rabanos*

1 Introduction

The relationship between law and coercion has always been a highly controversial topic in contemporary legal philosophy. After an initial phase in which there was a strong consensus on its essential importance for law¹, an apparent consensus on the exact opposite has emerged in the last decades². In recent years, however, several important publications have reignited the debate. They criticise the latter position and argue strongly in favour of considering coercion as a necessary or relevant property of law, as well as propose new ways to approach the debate and conceive the relevant question³.

The aim of this chapter is to contribute to the revitalised debate by arguing for the importance of coercion both for the existence of law and for the function(s) of law. I will argue that if law is seen as a tool or artefact, the fulfilment of its function(s) depends both on its existence and on its possession of a certain property that enables the achievement of that end. I will then show how coercion can be understood as necessary for the existence of law by analysing two main arguments for coercion-dependent existence: that coercion is motivationally necessary to ensure conformity to law, and that it is necessary to ensure the existence of a community. Finally, even if this argument fails, I will show that coercion can nevertheless be understood as the particular property that law has to fulfil its function(s). I will argue that this is the case for most of the possible functions of law, especially those related to behaviour-guidance and coordination-solving.

I will proceed as follows. In section 2, I will address some useful definitions of coercion and law and offer some insights into the view of law as an artefact and the functions ascribed to it. In section 3, I will set out some general theoretical positions on the relationship between coercion and law that will help to frame the content of the arguments that follow in the next section. In section 4, I will analyse the role and significance of coercion for the existence of legal systems (4.1) and for the function of legal systems (4.2). Finally, in section 5, I will offer some concluding remarks.

---

* J.A. Rabanos | julieta.rabanos@ius.bg.ac.rs | Postdoctoral Fellow in Theory and Philosophy of Law, University of Belgrade – Faculty of Law, bulevar kralja Aleksandra 67, 11000, Belgrade, Serbia.
2. On Coercion and Law: Some Definitions

2.1 About coercion

What exactly coercion is, or in other words, what the truth conditions are for statements such as “A coerces B to do X”, is still the subject of lively and ongoing debate\(^4\). When it comes to the relationship between coercion and law, the terms of the debate themselves are also subject to disagreement. It would not be possible to do justice to all these complexities and nuances here. However, a brief overview of some of them will be useful for the purposes of this paper, as they will help to establish the basic and general definition of coercion that will be used in sections 3 and 4.

In order to analyse the possible relationship between coercion and something else, it is first necessary to depart from some definitions. There are several ways to approach the definition of coercion. One way is to distinguish between means used by one subject to put another into a particular position. In this sense, MacCormick distinguishes between direct physical coercion and coercion by threat\(^5\). The former are cases where force is used to make a subject do X or submit to Z against their will and without real choice, while the latter are cases where threats are used to make a subject do X or submit to Z against their will and without real choice. MacCormick further distinguishes between coercion by direct threat and by indirect threat: the former involves the threat to bring about X by force, while the latter involves the threat to bring about Y if the subject fails to bring about X\(^6\). The core of MacCormick’s definition is that each of these means places the subject in a position in which they have “no real choice”. The main difference is that in one case there is only the application of force to make the subject to do X, whereas in the other case there is a “proposal” of sorts regarding the doing or not doing X\(^7\).

Following this latter approach, which involves a “proposal” or “rational compulsion”, two main families of accounts of coercion can be found: moralised and non-moralised accounts. Moralised accounts view coercion as a moral notion, where statements about coercion imply a moral assessment of some kind\(^8\), while non-moralised accounts reject this view. Within the moralised accounts, Alan Wertheimer’s “two-prong” account of coercion can be used as a starting point: “A coerces B to do X if and only if (1) A’s proposal creates a choice situation for B such that B has no reasonable alternative but do X and (2) is wrong of A to make such a proposal to B”\(^9\). The debate between the variants of moralised account centres on what counts as “reasonable alternative” and what counts as “wrong of A”.

Within the non-moralised accounts, three main approaches can be found: (1) the pressure theory, (2) non-moralised preference-baseline accounts and (3) a non-moralised coercer-intention account\(^10\). The pressure theory states that “A coerces B to do X” is to be understood as “A, by doing Y, has exert great psychological pressure on B to do X”. Coercion is purely a matter of psychological pressure, measured against B’s perception of A’s actions as exerting

\(^7\) Another name for these situations could be “physical compulsion” for the former, and “rational compulsion” for the latter. See v.gr. Lamond (2000), p. 49; Edmundson (2016), p. 464.
\(^8\) See v.gr. Oberdiek (1976).
pressure on their psyche. Non-moralised preference-baseline accounts have been offered to correct this heavily individualized index the pressure in pressure theory is measured against. One variant holds that “A coerces B to do X” should be read as “A, by doing Y, coerces B if and only if A has made B worse off compared to where B would be if Y had not been made”: B is worse off both when she does X and when she does not do X and suffers from A’s declared unilateral plan. Another variant holds that “A coerces B to do X” should be read as “A, by doing Y, coerces B if and only if B would prefer that A’s Y had never been done or made”. A third variant states that “A coerces B to do X” should be read as “A, by deliberately doing Y with the aim of bringing about an undesirable consequence for B should she fail to bring about X, has exert great psychological pressure on B to do X”.

Each of the above approaches has been the subject of extensive debate and criticism, which I will not go into here\textsuperscript{11}. However, there are two things are worth mentioning that arise from the definitional strategies we have briefly reviewed in the previous paragraphs. The first is that coercion can be defined by reference to both means and effects. If the definitional strategy follows the latter approach, coercion is generally understood as a success concept: there is no coercion in a situation where the effect (for example, the enormous psychological pressure) has not obtained. Failed coercion is therefore not coercion. If, on the other hand, the definitional strategy follows the former approach, then coercion is not a success concept: coercion can also exist in a situation in which the effect has not obtained, but some means were used, or a certain action or intention has obtained. In this sense, failed coercion—even if it is flawed— is still coercion.

The second thing worth mentioning is that a continuum of positions on coercion can be reconstructed, ranging from strict or narrow views to broad or wide views\textsuperscript{12}. On the one hand, strict or narrow views generally limit coercion to force, with the more extreme views considering only physical force\textsuperscript{13}. On the other hand, broad or wide views include elements other than force, such as arbitrariness\textsuperscript{14} or mere non-consented interference with the will\textsuperscript{15}. It is an open question whether these extremes can simply be replaced by “physical compulsion” and “rational compulsion”.

\section*{2.2. About law}

\subsection*{2.2.1. Law: functions and artifacts}

It is quite commonplace to think about law in terms of functions: it is commonly assumed that, as a somewhat byproduct of human creation and a social institution, it has been brought about with some specific purpose or to produce some specific effects. In this sense, law is seen as a

\begin{thebibliography}{9}
	\bibitem{11} A summary of them can be found v.gr. in Anderson (2011), but also in Edmunson (1995) and Lamond (2001).
	\bibitem{12} This is different from other classifications of strict and broad sense of coercion, classifying theories that use the concept. See Section 3, in fine.
	\bibitem{13} See v.gr. the classics Kelsen (1949), (1967); Hart (2012); Bentham (2010).
\end{thebibliography}
tool or a means to achieve a certain goal, value, or purpose. The disagreements have concentrated not on this commonplace idea per se, but on the details: what this function is, how to evaluate whether law fulfils it or not, whether intentionality is an essential component in determining this function, whether intentionality is an essential component in determining the existence of law itself, whether the intended function is something real or ascribed by the theorist, and so on.

One of the most important disagreements is related to the question of whether law is to be distinguished by its purpose or by the means it employs to achieve its purpose. This question can also be reformulated as the question of whether law is a functional or a modal kind. The former is distinguished by the particular purpose that it serves, while the latter is distinguished by the specific means it employs or provides to achieve particular ends or purposes. As Gardner warns, however, this disagreement is about a distinguishing criterion: to hold one position or the other is not to deny that law has some specific purpose or that law provides particular means to serve that end.

Among those who regard function as extremely important for a theoretical explanation of law, Ehrenberg has recently proposed four possible theoretical options: the views that (1) function is a necessary and sufficient condition for the presence of law; (2) function is a sufficient, though not necessary, condition for the presence of law; (3) function is a necessary, albeit not sufficient, condition for the presence of law; and (4) function is neither a sufficient nor necessary condition for the presence of law. To endorse this last position is to understand the function of law without committing to the idea that ‘it must always accomplish, or even aim at accomplishing, its function’ nor saying ‘that anything performing or aims to fulfil the function is law’. This becomes clear, according to Ehrenberg, if one understands law as a kind or genre of “institutionalised” abstract artefact.

An artefact can be understood as a product of intentional behaviour that has been intentionally modified in order to achieve (or enhance its ability to achieve) a certain goal and to communicate its identity or usability to a community of users. This identity is usually linked to the function of the artefact. However, even though functions are important to understand artefacts, they are not enough. Proper functions are the key to this understanding: they are the ‘characteristic ends that [something] yields, which explain its presence, development or selection’. In the case of artefacts, Ehrenberg says, proper functions are generally understood as “design functions” (rather than “use functions”): tasks that the creators or designers of the artefact envision when creating or designing the artefact. In this way, the fact that functions can be seen as value-laden is easy to recognise: ‘designers, creators, and users can all be understood to be interacting with the artifact for the reason that they believe it brings some

---

16 This original terminology is from Green (1996), later used by Gardner; see (2012), p. 207. Both authors seem to endorse the idea that the key for an understanding of law and its nature is to be found in its means, not in its ends. See Gardner (2012), p. 207-208, Green (1998), p. 123.
17 This explanation, in his words, includes the understanding of “the nature and normativity of law”. See Ehrenberg (2016), p. 180.
21 Ehrenberg (2016), p. 21. They have to be distinguished, Ehrenberg maintains, from “Cumming functions” (“benefits that a given trait or aspect contribute to a wider system of which it is part”); these functions have the issue that this wider system is not generally fixed, and it need to be fixed by an independent argument.
benefit or value to the world. Furthermore, the intentions of the creators or designers allow for consider both causes and effects and goals as part of the intended functions for the artefact. Finally, there is a clear distinction between “function” as a purpose or goal and “function” as “how it works”.

2.2. Some attributed functions

Famously, Raz proposed to analyse what he called the “social functions of the law”, distinguishing between direct and indirect functions, the former being divided into primary and secondary functions. In his words: «[d]irect functions are those the fulfilment of which is secured by the law being obeyed and applied. Indirect functions are those the fulfilment of which consists in attitudes, feelings, opinions, and modes of behaviour which are not obedience to laws or the application of laws, but which result from the knowledge of the existence of the laws or from compliance with and application of laws». Indirect functions are fulfilled as results of the laws’ existence and application of laws and their application in conjunction with other factors (such as the subject’s attitude towards the law and the existence of other social norms and institutions). Direct functions are not always fulfilled in this way, in the sense that obedience, for example, can be achieved without the subject’s knowledge of the existence of the particular law they are obeying.

Within the direct functions, Raz further distinguishes between primary functions and secondary functions. On the one hand, ‘primary functions are outward-looking, they affect the general population, and in them is to be found the reason and justification for the existence of the law’. There are four of them: (a) the prevention of undesirable behaviour and the securing of desirable behaviour; (b) providing facilities for private arrangements between individuals; (c) the provision of services and the redistribution of goods; and (d) the settlement of unresolved disputes. On the other hand, «secondary functions are the functions of the maintenance of the legal system. They make its existence and operation possible. They are to be judged by their success in facilitating the fulfilment of the primary functions by the law». There are two: (a) the determination of procedures for changing the law and (b) the regulation of the operation of law-applying organs.

Regarding the indirect social functions, Raz claims that «the social effects of the law which come under this heading almost always depend for their achievement on non-legal factors, especially the general attitude to the law and its interaction with social norms and institutions. Some of these functions are performed by particular legal institutions, others by the existence of the legal system itself».

---

23 Ehrenberg (2016), p. 27.
24 Raz (1979), p. 165 ff. For all the caveats he puts forward before attempting a classification, he claims that “[i]t seems to me that all legal systems necessarily perform, at least to a minimal degree, which I am unable to specify, social functions of all the types to be mentioned, and that these are all the main types of social functions they perform. These claims will not, however, be argued for here” (1979, p. 167).
In a recent book, Ehrenberg has analysed several theories dealing with the function of law and offered a general list of the theories his analysis encountered\(^{30}\). Unlike Raz’s proposal, this list is not comprehensive and non-exclusive, and deals with the main or characteristic functions that various authors have considered to be the primary explanatory functions of law. For our purposes, however, it may be of great theoretical interest.

If one follows Ehrenberg, then, among primary explanatory functions of law that appear in various theories these can be found: (1) justifying the state’s use or withholding of force; (2) implementing or enforcing the moral and political principles that a community holds dear; (3) general coordination and dispute settlement; (4) organising behaviour to solve coordination problems and to cooperate in the pursuit of the common good; (5) achieving a certain kind of order by subjecting people’s behaviour to the guidance of general rules; (6) provide guide to human behaviour and standards for criticising that behaviour; (7) creating new reasons for those subject to it through serving as a medium for communication of an intention to do so by those who author, implement, endorse, or ratify the law; (8) incentivizing behaviour in order to maximise economic efficiency; (9) providing an institutional framework for the specification, recognition and protection of contextual rights and duties by institutions in the broadest possible social environment.

Some comments related to this listing. At first glance, it looks as if some of these functions could be considered as instances of other, more general functions: (4) and (5) could be instances of (3), for example, and (7) and (8) could be instances of (6). In fact, most of these possibilities could just be sub-types of (3): as Ehrenberg himself says, ’[w]e saw that behaviour guidance, social ordering, predicting the state’s response to personal choices, communicating the society’s values, mitigating the demands of the individual upon others, and mitigating the demands of others upon the individual can all be harmonized within the idea that one primary function of law is to solve coordination problems’\(^{31}\). Similarly, some of these functions may be overinclusive: coordination and dispute settlement could be separated, depending on the specific definition of “coordination”. A more precise definition of “coordination” and “coordination problem” is necessary to assess or reject this inclusiveness.

Furthermore, some more classically ascribed functions seem to be missing here, such as the generic but well-known ‘maintenance of peace' or 'guaranteeing the security of the subjects’. At the same time, (1) could be subdivided into (1’) regulating the state’s use or withholding of force and (1'') justifying the state’s use or withholding of force. In this way, several theories that regard coercion as the specific content of the regulation of a legal system could be seen as promoting (1’), even if they do not deal with (1''). Finally, Ehrenberg points out a caveat: some of these purported 'ends' could well be seen as 'means'. Behaviour-guiding can be seen as a means to coordination-solving; coordination-solving can be seen as a means to achieve the common good; regulating the state’s use of force can be seen as a means to ‘maintaining the peace’; and so on. In this sense, drawing a line between ends and means and labelling something as an end or a means is difficult and ultimately depends on a theoretical choice.


---

\(^{30}\) See Ehrenberg (2016), ch. 8.

Within legal theory and philosophy of law, there are not only numerous questions about coercion and its relationship to law, but also several levels on which the discourse about it can be conducted. Indeed, different questions refer to different levels of analysis. Following Edmundson, we can at least between an explanatory and a normative level or sense of analysis. At the explanatory level, the concept of coercion is used to formulate questions relating to the way in which coercion can be seen as an element of law – be it ‘law’, ‘legal norms’ or ‘legal systems’. At the normative level, the concept of coercion is used to formulate questions about the legitimacy or justification of its presence or use by subjects and entities, such as the state or private individuals. Both levels of analysis and their associated questions may or may not be connected, and thus should be kept separate.

Looking more closely at these questions, coercion as an element or property of law can be used to answer all or some of the four main theoretical questions related to law: questions about identity (what is the criterion for an order or norm to be deemed ‘legal’?), efficacy or effectiveness (what is the criterion for an order or norm to be deemed ‘effective’?), efficiency (how can the law fulfil its function?), and normativity (how can the law guide human behaviour?). At this point, it should be noted that questions about coercion and “law” can indeed have different objects: They can relate to law, to legal systems or to (individual) laws. The answers to these questions can greatly differ.

Much of the contemporary discussion in legal philosophy about coercion and law, especially at an explanatory level of analysis, centres not on what exactly coercion is, but on whether coercion can be regarded as a necessary (at least in some sense) element of law. A continuum of positions can be reconstructed, ranging from law necessarily being coercion to law necessarily being not coercion. In the middle, several positions can be identified: among them, coercion being the distinctive feature of law qua social order (either its use or its regulation); coercion being the most prominent or salient feature of law (with respect to other social normative orders); coercion being just one of the paradigmatic features of law (not a necessary one), and law not (necessarily) needing or including coercion. Another way to categorise these positions would be to distinguish between those who claim that coercion is an essential element of law (either an instrumental essential element or a material essential element) and those who claim that coercion is not an essential element of law (either a contingent element or necessarily not an element).

Two very recent works have attempted to offer novel ways of clarifying the complexity of the debate and categorising the different theoretical positions. The first has both offered a new understanding of the ‘broad sense’ and the ‘strict sense’ of coercion and has proposed to

---

33 I have suggested this in Rabanos (2024b).
34 From now on, and from present purposes, I will use the terms ‘efficacy’ and ‘effectiveness’ as interchangeable synonyms.
35 In another venue (see Rabanos 2024b), I have associated this question with the question about the criterion to deem a legal order or norm existent. However, this association without any caveats can be challenged; see Section 2, in fine.
37 The latter holds particularly true among those who, in some sense, consider the issue as part of a wider discussion about the characteristics of conceptual analysis and the methodology of concept-formation.
38 I am basing this reconstruction on Lamond (2001) and Bobbio (1965), (2014).
39 For instance, positions using moralised baselines to identify cases of coercion and considering law to be always morally wrong in the use of force or the restriction of choice.
40 For instance, positions using moralised baselines to identify cases of coercion and considering law to be (at least prima facie) morally right in the use of force or the restriction of choice.
distinguish between different senses of 'dependence' of law on coercion\(^{41}\). Regarding the former, Miotto has proposed to classify theories according to the importance (or lack thereof) they attach to the development of comprehensive theories of coercion. “Strict sense” theories maintain that a complete theory of coercion is necessary to answer some of the questions of a theory of law, whereas “broad sense” theories do not concern themselves with the details of coercion and use the concept in a broad and vague way\(^{42}\). Regarding the latter, Miotto has proposed to distinguish between five senses of “dependence”: ontological, epistemic, analytical, conceptual, and pragmatic\(^{43}\). Thus, “\(a\) depends on \(b\)” can be read as expressing: “\(a\) exists in virtue of \(b\)” (ontological dependence), “our ordinary concept of \(a\) involves or makes reference to \(b\)” (conceptual dependence), or “\(a\) properly works in virtue of \(b\)” (pragmatic dependence). These senses of “dependence”, according to Miotto, are at least \textit{prima facie} independent.

The second work has proposed a new taxonomy for theories on the jurisprudential importance of coercion and also to change the relevant question when analysing the debate from a 'modal question' to a 'metric question'\(^{44}\). Regarding the former, Gkouvas has proposed to base this new taxonomy on the idea that the truth of judgements about the necessity or contingency of coercion for a system of governance by law depends on a practical standpoint: either a deontic perspective (an agent who considers law as a distinct source of duties and rights) or an axiological perspective (an agent who considers law as a valuable instrument for promoting certain values that exist independently of the system)\(^{45}\). Regarding the latter, Gkouvas characterises the question that the contemporary debate seeks to answer as a 'modal question': the question of whether governance by law is necessarily or contingently coercive. He argues that 'necessity', the key element of this question, can be conceptualised or described, but cannot be measured or used\(^{46}\). For this reason, he proposes to change the relevant question to the 'metric question': a question about how (much) legally inexorable the truths about the legal relevance of situations of state coercion are\(^{47}\).

\section*{4. On Coercion and Law: Functions and Existence}

In sections 2 and 3 we have seen an overview of the different approaches to the concepts of coercion and law, as well as a number of frameworks for considering their general relationship. In this section, some of these approaches and frameworks will be adopted and assumed. On the one hand, I will endorse the view that law is a kind of artifact with a particular purpose or end, and that it is not necessary to discuss either the status or origin of this purpose or end nor the precise nature of the artifact law is\(^{48}\). On the other hand, I will endorse a general, primarily non-moralised\(^{49}\) approach to coercion, making space for both effect-related and means-related

\begin{footnotesize}
\footnotesize
\begin{enumerate}
\item See Miotto (2020).
\item See Miotto (2021), p. 80-1.
\item See Miotto (2021), p. 75 ff.
\item See Gkouvas (2022).
\item Gkouvas (2022), p. 8.
\item Gkouvas (2022), p.1.
\item Gkouvas (2022), p. 21 ff.
\item For instance, see Khalidi and Murphy (2021).
\item Even if, in my opinion, much of what will be argued for in this section is also compatible with adopting a moralised approach.
\end{enumerate}
\end{footnotesize}
approaches\textsuperscript{50}. Finally, the following arguments operate at an explanatory rather than a normative level\textsuperscript{51}; they relate primarily to questions of effectiveness and efficiency, not identity or normativity\textsuperscript{52}; and they relate primarily to law in the sense of legal systems, not laws or legal norms.

Thus, if law is regarded as an artefact, it necessarily has some kind of purpose or function(s); and – I will further assume – the fulfilment of its function(s) depends both on its existence and on its possession of a certain property that enables the achievement of that purpose. In other words, for law to be able to fulfil its purpose or function, two conditions must be met: (1) law must exist (pre-requirement); and (2) law must have some feature that enables it (in some way) to fulfil the purpose or function (requirement). Whether this dependence or necessity can be seen as ontological, metaphysical, conceptual, or pragmatic remains an open question to be addressed.

In the following points, I will first show how coercion can be understood as necessary for the existence of law by analysing two main arguments for coercion-dependent existence: that coercion is motivationally necessary to ensure conformity to law, and that it is necessary to guarantee the existence of a community (4.1). Even if this argument is not convincing, I will show that coercion can still be understood as the particular property that law has in order to fulfil its function(s). I will argue that this is the case for most of the possible functions of law, especially those related to behaviour-guidance and coordination-solving (4.2).

4.1. Coercion and the Existence of Law

In order to examine the possible dependence of law on coercion, I propose to depart from a commonplace idea: that a legal system, in order to exist, must be efficacious. This effectiveness or efficacy can in turn be seen as dependent on two elements: (i) the existence of a community and (ii) a sufficient degree of general conformity with the bulk of norms of the system. One could add a third element: (iii) specific compliance with some of the system's norms by specific members (so, if rules change, they also change behaviour)\textsuperscript{53}.

There are two classic arguments in favour of coercion-dependent efficacy\textsuperscript{54}. The first is that coercion is necessary to ensure (ii), i.e., a sufficient degree of general conformity. There are two ways of understanding this argument: an extreme reading and a moderate reading. The former claims that coercion is a necessary and sufficient condition to ensure (ii): not only needed, but also sufficient in itself. The latter claims that coercion is only a necessary condition:

\textsuperscript{50} This also means, albeit maybe not directly, that there would also be space for coercer-centred and coercee-centred approaches.
\textsuperscript{51} See section 2 of the present chapter.
\textsuperscript{52} See section 3 of the present chapter.
\textsuperscript{53} It is worth noting here that authors such as H.L.A. Hart, even considering both (ii) and (iii) normally important, has argued that in an extreme case of legal system it can be considered as efficacious (and existent) just by (iii), even if (ii) does not obtain. See Hart (2012), pp. 116-117. This is so because, in his view, it is not possible for officials to merely “conform”: they have to adopt the internal point of view, even if no other member of society does. Kelsen can be also considered to ultimately focus on official behaviour to determine the effectiveness of a legal system: for him, a legal system is effective when the socially undesirable behaviour present in the condition of application of its norms is by and large not committed or, if committed, the corresponding sanction is by and large determined and executed. See Kelsen (1991), pp. 138. However, in my opinion, the behaviour of both officials and non-officials would for Kelsen be part of (ii).
\textsuperscript{54} See Lamond (2001), p. 45.
something that, even if needed, does not suffice in itself. In turn, the second argument is that coercion is necessary to limit the socially destructive levels of violence within a social group and thus enable the continued existence of a community (i).

These arguments seem to be as commonplace as the effectiveness-dependent existence idea. On the one hand, they are grounded in generalisations about human nature and the nature of human social life. These generalisations are compatible with a whole range of positions, ranging from a Hobbesian pessimistic approach to a Rousseauian good-savage view, and tend to coincide in considering real or potential problems related to resources. On the other hand, they are modest in two respects. The first is that, at least on a moderate reading, they do not claim that every law must be coercive or that coercion is the primary motivation for general conformity to law or the primary means of controlling violence. Rather, they claim that they ensure (i) and (ii). The second way is that, at least on the least compromised reading, they show only that human law is coercive in this sense. This is because of the particular characteristics of the subjects to whom it applies, and because it might not be of interest to concern ourselves with other kinds of potential (and non-existent) subjects.

Although these arguments are widely commonplace, they can face different challenges: (1) regarding the generalisations that underlie them; (2) regarding the kind of coercion they assume; (3) regarding the kind of necessity or dependence they establish; and (4) regarding the assumption of a connection between effectiveness and existence. Let us briefly consider these challenges and assess their soundness.

(1) The generalisations about human nature and the nature of human social life that underlie these arguments can be challenged both in themselves and challenged regarding their conclusiveness to prove that coercion is a necessary condition for (i) and (ii). For example, Lamond points out that these generalisations widely overlook other means of ensuring effectiveness, such as socialisation, social pressure to conform, benefits to conformity (which cannot be achieved without cooperation), and acceptance of the authority of the law. He further argues that they are too cavalier in assuming that coercion must play a central role in the effectiveness of legal systems: even if no society would dispense of coercive measures, this does not show that coercion is necessary to maintain effectiveness; it would be enough for just one legal system to remain efficacious without the use of coercion to prove otherwise.

Interestingly, this last argument could also be turned around and the burden of proof reversed: The fact that no known society has dispensed of coercive measures, that is, no known legal system has remained efficacious without the use of coercion, might actually show that coercion is necessary to maintain effectiveness. In the face of that knowledge, it might indeed be reasonable for the burden of falsification to be placed on those arguing for the possibility of the unknown, rather than the other way round. In any case, the “overlooking other means of ensuring effectiveness” is a double-edged sword. On the one hand, socialisation and education can be seen as devices for internalising coercion. On the other hand, cooperation (and its benefits) can only be achieved, at least in some cases, if potential participants are aware that there are measures in place to deal with free riders and non-participants. Finally, this criticism seems directed to the extreme reading of the main argument (coercion as a necessary and sufficient condition); it would lose force in the face of the moderate reading.

---

55 See v.gr. Olivecrona (1939), pp. 140 ff, especially 141-142, when he develops the argument about the indirect effects of coercion.

56 I will develop this in section 4.2, in fine, of the present chapter.
(2) Even if the necessity of coercive measures for effectiveness could be demonstrated by these arguments, they would still not prove that these measures must be legal ones. In this case, Lamond argues, they are perfectly compatible with the existence of non-legal coercive measures that are recognised or authorised by the legal system but are not part of the legal system. This would, of course, mean that the dependency link between coercion and law remains, but rather than being an aspect of the nature of law, coercion would more properly be seen as an aspect of the nature of social life. In any case, the arguments would be useful to demonstrate, at best, that legal system *qua* social standards (which regulate basic aspects of the social group’s life and maintain it) require coercive support. The degree of coercion required would depend on the strength of the non-legal sources of conformity.

In my view, this critique seems to assume that coercion is seen as an identity criterion for law (i.e., a criterion for distinguishing law from other entities); but this need not be the case. If law is a part of social life or a subspecies of social practice, then it would reasonably follow that law shares the characteristics of social practices – and this would indeed include coercion. The identity criterion may in any case be the kind of coercion that is used. In any case, there is a very interesting point concerning the last part of this critique. Lamond points out that the degree of coercion would definitely be higher in pluralistic societies. The fact that societies are pluralistic and do not consist of the same kind of people might be an argument in favour of the dependence of law on coercion.

(3) At best, they make an argument of “natural necessity” or “instrumental necessity”; or, to use a different terminology, they make an argument of pragmatic dependence. As such, these arguments would say nothing about the nature of law or the concept of law, since they are not concerned with necessary properties or with establishing conceptual or logical truths. They would not be able to demonstrate that coercion is a constitutive property of law, which is made clear by their care to refer to human law and by the fact that we can conceive a possible world in which a phenomenon recognisable as law exists, but no trace of coercion is present.

In my view, the success of this critique depends on the endorsement of both a particular way of understanding conceptual analysis (linked to a kind of metaphysical existentialism or otherwise accepting the existence of ‘natures’) and a particular law to understand law as an artefact with a particular function (the generic allocation of rights and imposition of duties). The rejection of one or both endorsements can significantly weaken the strength of the critique. In the first case, it is possible to reject this understanding of conceptual analysis by, for example, endorsing a paradigmatic or central case methodology, including those that treat the central case as an ideal type or those who construct the central or paradigmatic case as containing all possible properties with respect to a word, or a family-resemblance or cluster methodology, where different instances of a term share only some properties and not others, but there is neither a paradigmatic case as defined above nor a closed set of all possible properties. In the second case, it is possible to reject this understanding of the function of law, for example by viewing the function to be about guiding behaviour or solving coordination problems at a motivational level of analysis, or even by taking other positions, such as that the function is the justification of the state’s use of force.

In my view, the success of this critique depends on the endorsement of both a particular way of understanding conceptual analysis (linked to a kind of metaphysical existentialism or

---

57 This is, for example, Kelsen’s position. See Kelsen (1957).
58 I will touch upon this in more depth in sections 4.2.1. and 4.2.2 of the present chapter.
otherwise accepting the existence of 'natures') and a particular way to understand law as an artefact with a particular function (the generic allocation of rights and imposition of duties). The rejection of one or both endorsements can significantly weaken the strength of the critique. In the first case, it is possible to reject this understanding of conceptual analysis by, for example, endorsing a paradigmatic or central case methodology, including those that treat the central case as an ideal type or those who construct the central or paradigmatic case as containing all possible properties with respect to a word, or endorsing a family-resemblance or cluster methodology, where different instances of a term share only some properties and not others, but there is neither a paradigmatic case as defined above nor a closed set of all possible properties. In the second case, it is possible to reject this understanding of the function of law, for instance, by conceiving the function to be about guiding behaviour or solving coordination problems at a motivational level of analysis, or even by taking other positions, such as that the function is the justification of the state’s use of force.

In any case, as we will see in the next point, considerations related to conceptual dependence and pragmatic dependence seem to overlap, if not collapse, when conceiving law as an artefact (and, specifically, a functional kind). If this is so, then the argument, even if it is about pragmatic dependence, could also be regarded as an argument about conceptual dependence. What is more: if this kind of conceptual analysis is accepted and the presupposed nature of objects can be related to the conditions for their existence, then it would be also related to an argument about ontological dependence. Whether conceptual analysis is about conceptual or metaphysical dependence is, of course, a question that I will leave open here.

(4) The assumption that existence and effectiveness of law are connected and that what is needed for effectiveness is also needed for existence can be contested in two ways. The first is to deny that sufficient general conformity (i.e., ii) is a product of coercion (or the coerciveness of the system). The second is to deny the argument that legal systems exist only if they are efficacious. This would mean rejecting a connection between ontological dependence and pragmatic dependence: that is, rejecting the claim that law exists in virtue (at least partially) of coercion if law properly works in virtue of coercion.

We have already seen one way to assess the first challenge, which is to reject the extreme reading of the claim (and to recognise other sources of conformity) and to consider it at least plausible as long as no empirical analysis falsifies it. Let us now turn to the second challenge, which may be a more novel critique. Miotto points out that ‘whether a legal system owes its existence to facts related to coercion and whether a legal system owes its efficacy to facts related to coercion seem to be fairly independent questions’ and that therefore ‘claims about legal system’s ontological dependence on coercion (…) do not follow automatically from claims about pragmatic dependence’. Anyone attempting to make the leap from pragmatic to ontological dependence must put forward the arguments for doing so and not simply take them for granted.

This is a fair point. I think no one would deny that all assertions in the context of an academic debate need to be argued for, and I also agree that in some cases such claims are taken as axioms rather than theses to sustain. However, in my view are two possible arguments to
respond to this challenge. The first is that although questions about the conditions for existence and questions about the conditions for efficacy may be theoretically independent, they also may overlap or be related under certain circumstances. One of these is when practices and behaviour are central to the analysis. In this sense, if X exists because (it consists of certain practices or behaviours, and these practices or behaviours cannot function or exist without Z, then it would follow that X cannot exist without Z. The second argument is related to the consideration of law as an artefact (and a functional kind) and the possibility that questions about pragmatic, conceptual and ontological dependence may strongly overlap or directly collapse. We will analyse this in the following section.

### 4.2. Coercion and the Functions of Law

Let us now turn to analyse a second way of exploring the possible dependence of law on coercion: that coercion is a necessary feature for law in order to function (in order to achieve a particular purpose or goal). This line of reasoning can arise, among other things, from the rejection of the first way of exploring the relationship, from a completely unrelated line of exploration, or from some of the last suggestions we saw at the end of section 4.1.

As we saw in section 2, there is disagreement among proponents of the idea that law has a kind of function there is disagreement whether law should be considered a modal kind (identified by its means of achieving the function) or a functional kind (identified by its function itself). This is significant insofar as doubts might arise as to whether some of these arguments concerning coercion and function treat coercion in a purely modal or in a purely functional way. In any case, I will try to show that this makes no real difference here and that the following discussion can be useful both for those who understand law as a modal kind and for those who understand law as a functional kind.

A complete analysis of all possible functions of law and the role that coercion might play in each of them might prove to be an impossible task, especially given the heterogeneity and diversity of possible functions. However, I will follow here the idea that a large part of them can be seen as specifications of a broader co-operative function: I will thus deal with the function of “guiding behaviour in a wide sense” and the function of “coordination in a strict sense”. Finally, in selecting the functions to be analysed, I will not deal with those directly related to the regulation or justification of coercion, since the dependence they create –even if its type could be debated– seems pretty clear.

#### 4.2.1. Guiding behaviour in a wide sense

When thinking about the functions of law, one widespread contemporary commonplace idea is that this function revolves around ‘guiding behaviour’66. This idea can be expressed in different ways, such as ‘providing guidance’ and ‘providing reasons for action’, or with qualifying statements such as ‘subjecting human behaviour to the guidance of general rules’. What underpins these different approaches is the common assumption that law provides people with some patterns of behaviour that should be observed or conformed with, and that it does so in...
such a way that people can both recognise and know these patterns and know that other people also have access to this recognition and knowledge.

It can be assumed that this behaviour-guiding function always provide for coordination and coordination-solving in a wide sense. As such, it can be argued that law, as it provides guidance at the same time to a group of people in a variety of situations, then always performs a coordination function – even when that behaviour-guidance is directed at actions or circumstances involving only one person or a situation where no clear coordination problem is in sight, or at a particular action where coordination with other participants is not required. However, for present purposes, I will admit the possibility that the behaviour-guiding function does not necessarily amount to coordination in the strict sense. I will examine the latter in the next point.

Would coercion be a necessary feature for law to “guide behaviour” in this sense? This question could be answered from a micro- and a macro- perspective, i.e., from an individual and a collective perspective, and considering complex and heterogeneous societies on the one hand and simple and homogeneous societies on the other. I will start here from a macro perspective, which includes considerations from a micro perspective.

Simple and homogeneous societies would, by definition, consist only of a certain kind of people with a certain attitude towards the law. For our purposes, I propose to consider three societies consisting of three ideal types of people: “bad man” type people, law-abiding or committed people, and puzzled people. The “bad man” society appears to be quite straightforward: coercion seems to be needed for any attempt of behaviour-guiding regarding non-law-abiding and purely self-interested people. Indeed, this scenario is the one that those who want to reject the dependence of law on coercion tend to focus on in their arguments and criticisms. In my opinion, the real challenge is posed by the other two societies.

The “law-abiding” society has been, in fact, notoriously used in arguments against the thesis of the necessary connection between coercion and law. By definition, it consists of people who are so inclined to be guided that they only need knowledge of the relevant behavioural patterns in order to be guided by them; nothing more is needed. From a micro-perspective, this

---

67 I will not delve here on the discussion about the Rule of Law and its requirements, which would merit an entire analysis in itself. However, I have generalised here some of the most commonly accepted requirements among those who endorse a thin Rule of Law conception, as they are very basic elements underlying the idea of “guidance”. For a development, see v.gr. Celano (2013).

68 See section 4.2 of the present chapter.

69 My usage of “complex” and “simple” does neither include nor carry any value-judgement regarding these societies. The choice has been made purely to better represent the fact that one set is composed not only by a multiplicity of elements, but also the elements are different from one another, while the other is not.

70 Here, we can also add a third thread of analysis: whether we are considering efficiency in guidance or not. In the latter, an argument about concentration in just public communication features can be made. In the former, considerations about “effective” guidance need to me make as just public communication may not be enough. I will not consider this complexity here.

71 Bezemek proposed, on the basis of Schauer (2015), the possible existence of a multiplicity of kinds of people in regard to a multiplicity of possible attitudes towards the law. He recognises a ‘bad man’, a ‘good man’, a ‘good citizen’, a ‘puzzled man’ (or ‘ignorant man’), a ‘moral man’, a ‘moral person’, and “tribes” such as the ‘Society of Angels’ and the ‘Race of Devils’. See (2016), pp. 15-16). Here, I consider only a simplified version of the possible types of people, as I think the point I am making can come across without more complexity. However, in my opinion, Bezemek’s taxonomy remains of great interest for further analysis.


73 Based on some arguments about the “society of angels”. See v.gr. Raz (2009), 149 ff.


could actually mean that the law could very well fulfil its purpose of behaviour-guiding without having to resort to or exercise coercion. From a macro perspective, however, some doubt can be shed to that conclusion. It is possible that the law-abiding attitude that characterises the members of this “law-abiding” society is wholly or partly related to their knowledge or belief that the other members of society are equally law-abiding. If for some reason this knowledge is not widespread or if it is not possible for them to form the relevant belief in a sustained, unequivocal, and lasting way, then the initial certainty that the law will fulfil its purpose without recourse to anything else becomes a mere probability or even just a possibility. And this is where coercion comes into play again: its persistent presence can serve as a foundation on which “law-abiding” members can build their (even if false) belief in the “law-abiding” attitude of others and in this way maintain their own.

Finally, the “puzzled” society is an interesting case to analyse. By definition, a “puzzled” society would be very similar to a “law-abiding” society, in the sense that it consists of people who want to know what is required but do not know what in fact is required or that have the disposition to comply (not to merely conform) the law for the only reason that it is the law. In fact, the only difference between a “puzzled” person and a “law-abiding” person might be that the latter already knows what is required, while the former lacks that knowledge. Another difference might be that in the second definition, the “puzzled” person wants to know what is legally required in order to comply with it, while it is an open question whether the other “puzzled” person is concerned with what is legally required or what is morally required (which may include knowledge of what is legally required). In either case, the considerations at the end of the previous paragraph would also apply to the case of the “puzzled” society. They would also include the question regarding the way other members of society would access and interpret the content of “what is required”, and whether all members would access and interpret the same thing.

In contrast, complex and heterogeneous societies are by definition made up of a diversity of people with different backgrounds, values, goals, and motivations. This means that there is a multitude of people whose reasons for accepting or rejecting law’s guidance are potentially divergent, even contradictory. In other words, these societies are made up of a combination of “bad man” type people, “law-abiding” people and “puzzled” people in varying proportions on the one hand, and people who can potentially fluctuate between different attitudes towards the law depending on various factors (including time, circumstances, specific content of certain norms, etc.) on the other. In this context, it seems difficult for a member of such a society to be certain about others’ attitudes towards the law or to form certain and lasting beliefs about it. We have already discussed what can follow from such a situation and what role coercion might

76 It is worth noting that some authors, on the basis of recent findings by evolutionary psychologists, have suggested that the law-abiding attitude in human beings is some kind of adaptative mechanism for survival. According to this view, human beings would have a disposition to cooperate in order to survive, and this disposition is manifested in the law-abiding attitude. I thank Jorge Silva Sampaio for letting me know about this point through his recent unpublished research.

77 This is Hart’s definition. See v.gr. Hart (2012), pp. 40-41.

78 This is Schauer’s definition. See v.gr. Schauer (2015), p. 42.

79 Note that here I am not talking about mere conformity, but something more. We have already seen, in the previous point, arguments related to the diversity of non-legal sources for conformity with the law and that this mere conformity is perfectly compatible with the fact that a subject might not know the content of the law (for instance, it is perfectly possible that someone does never kill another human being because of their moral code, without any knowledge of the criminal law provision prohibiting). However, to “be guided” by law, a subject must at least know the existence of the pattern of conduct and that knowledge must be part of the reason for their action or omission.
play in this. Furthermore, if this diversity is to be respected without forcing a kind of homogenisation (e.g., through education), then the presence of coercion seems to be relevant as an all-round neutral motivating factor that applies – albeit perhaps to varying degrees\(^{80}\) – to all kinds of people.

4.2.2. Coordination in strict sense

Let us now turn to the idea that the function of law is coordination in a strict sense. As anticipated, for the present purposes this function is understood as a function of strictly solving coordination problems as opposed to just guiding behaviour in a wide sense. In order to analyse whether coercion is a necessary feature for law in order to “solve coordination problems”, it is crucial to start with a definition of “coordination problem”.

There are several ways to define what a coordination problem is or what characteristics would make a situation a coordination problem. A first rough definition could be as follows: A coordination problem exists when a group of individuals has a reason to achieve a certain outcome \(R\), the achievement of which depends on their joint action, and: 1) there are different alternatives for joint action, each of which is sufficient to achieve \(R\); and 2) \(R\) is achieved if all individuals agree on a certain strategy (whatever that may be). We can already see here that there can be two types of coordination problems: subjective coordination problems and objective coordination problems. In the first case, the reason for achieving the outcome \(R\) is a subjective reason that depends on each individual: for example, their preferences or desires. In contrast, in the second case, the reason for achieving the outcome \(R\) is an objective reason: the outcome \(R\) (for example, the “common good”) is such that all individuals have an objective reason to achieve it, regardless of their subjective judgements.

This initial definition can be refined somewhat, and two different approaches can be found: a ‘traditional approach’ and a ‘non-traditional approach’. Starting from the basic assumption that the interests of the actors involved are aligned or converging, a traditional approach defines a coordination problem as a situation in which: 1) A group of agents must choose between mutually exclusive alternative courses of action; 2) The direct consequential outcomes depend both on their own choices and on those of others; 3) Each agent can fully, transitively, reflexively (and unconditionally) rank outcomes in an order of preference and seeks to do their best to respect that order; 4) Strategic choices are always conditional on the expectations of others' choices; 5) All agents share the knowledge of 4) (and 3)); and 6) There are at least two appropriate coordination equilibria (i.e., choices for all that are such that at least one of them would be worse off if another agent had chosen differently)\(^{81}\). In contrast, a non-traditional approach (such as that proposed by Finnis) defines a coordination problem as a situation in which: 1) significantly beneficial outcomes (that would otherwise be practically impossible to achieve) are achieved by a significant number of people through coordinated action; 2) there is sufficient shared interest to make such coordination attractive; and 3) The problem is to choose an appropriate pattern of coordination so that coordination can actually take place\(^{82}\).

\(^{80}\) This difference in degree might be caused by inequalities: even with the same coercive measures in place, people might experiment different degrees of affectation depending on personal characteristics (gender, race, social class, and so on) or contextual circumstances.


\(^{82}\) See Finnis (1984), p. 133.
I cannot undertake a complete and in-depth analysis or comparison of these two approaches here. For the present purpose, however, it is interesting to note that at least four main differences can be found between these approaches: (i) the traditional approach assumes transitivity in the ranking of outcomes, while the non-traditional approach rejects this; (ii) the traditional approach assumes that agents are guided only by instrumental rationality with the sole aim of maximising their preferences and interests, while the non-traditional approach rejects this and postulates that agents are guided by a different kind of rationality with a different aim; (iii) the traditional approach is compatible with both subjective and objective coordination problems, while the non-traditional approach is only compatible with objective coordination problems; and (iv) the traditional approach assumes that there is a convergence of interests between the agents involved and therefore cannot be applied to a situation where there is no such convergence; the non-traditional approach assumes that there is always a convergence of interests and that any non-convergence is only apparent, and can therefore be applied to virtually any situation of any group.

Would coercion be a necessary feature for law in order to “solve coordination problems”, thus defined? Roughly speaking, one could argue that the only necessary feature for law to do that would be to exist and to have some capacity to give salience to a coordination pattern and to communicate or signal it publicly and widely to all relevant participants. This simple public salience might be sufficient for participants to form mutual expectations about each other’s choices and to act on the basis of those expectations. However, it might not be sufficient: not in a context of high uncertainty about the reasons the other participants might have for choosing or not choosing the same coordination equilibrium; not in a context of uncertainty in which the goal is only achieved by the joint action of all members of a group, but there are some of them who might not be interested in achieving the goal; and not in a context in which, even if the members share the interest in achieving the goal, maximising their individual interest would constitute a free-rider situation (in which everyone else acts together but they do not act) and this might jeopardise the achievement of the goal.

In these cases, simple public salience by itself would not suffice, since the agents cannot be sure whether the other agents will also consider this coordination pattern as salient and act accordingly. Thus, it seems that something else is required for the fulfilment of the function of coordination-solving. This missing feature might be coercion: the capacity to impose coercive measures in case the participants do not choose the signalled option, thus giving it a qualified salience. This capacity is fundamental for any agent when deliberating on what to do: they will consider that a certain solution has acquired a salient character given that a coordination pattern has been chosen and signalled, accompanied by the threat of coercive measures in case of non-conformity. Why is this the case? Because the agent knows that the coordination problem cannot be solved if there are at least some subjects who do not conform their behaviour to a single course of action (either because they do not share the reasons or because they do not know them). A coordination pattern that publicly sets out a course of action and backs it up with the threat of coercive measures becomes the salient option for each agent in three ways. First, it gives all agents a strong prudential reason (the threat of coercive measures) to

---

83 I explored this in Rabanos (2024a), chs. 5 and 6.
84 This takes us back to section 4.1 of the present chapter, and the analysis of coercion and the existence of law.
85 It might be enough because this salience modifies the factual context, altering what would be “reasonable” to do according to the balance of reasons. See Green (1988), pp. 111 ff. However, this depends on both the formation of the general mutual expectation and the effective convergence of interest regarding the achievement of the goal.
participate in the coordination and to conform their behaviour to this coordination. Second, it gives each agent reason to believe that any other agent who desires coordination will also consider it as the salient option for the same reasons. Third, it gives each agent reason to believe that both free-riders and those who are not interested in coordination (but whose actions may be necessary to achieve the goal) will be forced to participate and to conform.

5. Concluding remarks

Undoubtedly, the current revitalised debate on the relationship between law and coercion will prove extremely fruitful, as both commonplace and novel ideas are put forward from very different theoretical standpoints and the analysis is linked to insights from a variety of other fields of study. Within this framework, I have proposed here to follow the path of those who have taken a step to the side from the traditional main ways of framing the discussion and to analyse the significance of coercion both for the existence of legal systems and for the functions of legal systems, departing from the idea of taking seriously the idea that law is an artefact. One of the main conclusions is that the space for the necessity of coercion would still be present when considering the behaviour-guiding and coordination-solving functions of legal systems even in “law-abiding” and “puzzled” societies. It remains an open question for future research endeavours whether this can also be said of any other function that can be attributed to law.

Acknowledgments: This essay was presented at the Special Workshop ‘Sanctions: An Essential Element of Law?’, organised by Christoph Bezemek, Nicoletta Lavadac Bersier, and Frederick Schauer, at the IVR World Congress 2022 in Bucharest (Romania). It was then presented at the Belgrade Legal Theory Group, established by the IVR Serbia (Serbian Association for Legal Theory and Philosophy), in 2023. I would like to thank the audience and the other participants of both events for their most helpful and challenging comments, that contributed to substantively improve the ideas contained in this chapter.

This paper results from research conducted within the Horizon Twinning project “Advancing cooperation on The Foundations of Law – ALF” (project no. 101079177). The project is financed by the European Union.

References


Olivecrona K (1939) Law as Fact (1 ed.). E. Munksgaard, Copenhagen.


