THE EUROPEAN ACADEMY OF LEGAL THEORY SERIES

General Editors
Professor Mark Van Hoecke
Professor François Ost
Professor Luc Wintgens

Other titles in this Series

Moral Conflict and Legal Reasoning
Scott Veitch

The Harmonisation of European Private Law
Mark van Hoecke and *François Ost* (eds)

Law as Communication
Mark van Hoecke

Legisprudence
A New Theoretical Approach to Legislation
Luc Wintgens (ed)

On Law and Legal Reasoning

FERNANDO ATRIA

HART PUBLISHING

OXFORD – PORTLAND OREGON
2001

1

Constitutive Rules, Institutions and the Weightier Matters of the Law

> Football is not a matter of life and death:
> it's much more important than that.
> BILL SHANKLY

This book is intended as a contribution to the literature on legal theory and legal reasoning. In particular, it seeks to examine the relations that obtain between law and a theory of law and legal reasoning, and a theory of legal reasoning. Two features of law and legal reasoning will be particularly important in this regard: law is *institutional*, and legal reasoning is *formal*. These two features are so closely connected that it is reasonable to believe that in fact they are simply two ways of looking at the same issue. I hope this will become clearer as the focus of the book shifts from the institutional nature of law, with which this chapter is concerned, to the consequences of this for legal reasoning, which will entertain us in the following chapters.

The word "institution" encompasses a wide range of ideas that, at best, bear a tenuous family resemblance to one another. As a form of legal literature, it has a long and venerable tradition, going back to the *Institutes*, or brief expositions of the law, common since Roman times. In the literature of speech acts, the analysis of the so-called "institutional concepts" such as promises and the like plays a crucial role. A related legal usage is that in which a "trust" is an "institution" peculiar to the *common law*.

I believe that the idea of "institutional" facts, the existence, consequences and termination of which depend upon the existence and application of rules and the occurrence of some brute facts, can be highly successful in dealing with some of the insights offered by the literature on legal reasoning. To use this idea in understanding the law *and* legal reasoning, however, a distinction between "brute" and "institutional" facts is not enough. All institutional facts might be equal, but some of them are certainly more equal than others.

TWO CONCEPTS OF RULES

In his seminal article "Two concepts of rules", John Rawls drew a distinction between what he called the "summary" and the "practice" conception of rules. For Rawls, the "summary conception" regards rules as summaries of previous

decisions. Each of these is made on the balance of all relevant considerations, and some kinds of cases will be recurrent: "thus it will happen that in cases of certain kinds the same decision will be made either by the same person at different times or by different persons at the same time" (Rawls, 1955: 34). In this context, a rule is formulated to encompass cases of that kind, so the next time the agent will have no need to ponder all the applicable moral considerations, and will instead be able "simply" to apply the rule. We see why in this conception rules are actually "summaries" of previous decisions: "rules are regarded as reports that cases of a certain sort have been found on *other* grounds to be properly decided in a certain way" (Rawls, 1955: 34).

The "summary conception" of rules is contrasted by Rawls to what he calls the "practice conception", in which rules are seen as

> defining a practice. Practices are set up for various reasons, but one of them is that in many areas of conduct each person's deciding what to do on utilitarian grounds case by case leads to confusion, and that the attempt to coordinate behavior by trying to foresee how others will act is bound to fail. As an alternative one realizes that what is required is the establishment of a practice, the specification of a new form of activity; and from this one sees that a practice necessarily involves the abdication of full liberty to act on utilitarian and prudential grounds. It is the mark of a practice that being taught how to engage in it involves being instructed in the rules which define it, and that appeal is made to those rules to correct the behavior of those engaged in it (Rawls, 1955: 36).

In this chapter the related concepts of "practice", "institution" and "rules" are to be discussed. In that discussion, I will look at games and the law as providing the standard instances of practices. Games, in fact, will constitute one paradigmatic case. This is not because games are the most important practices: indeed they are not. But they do present the features of institutional facts in a particularly pure and clear-cut way. They can be sliced off from their environment and studied in ways in which other institutional facts cannot. This is an important point in itself, and we shall see later on that an explanation for this feature of games is called for. For the time being I just want to emphasise that my choosing games as one paradigm of practices is intended in a purely analytical sense. Games are a kind of luxury we get on top of what is really important, rather than the real thing. But this need not bother us. To claim paradigmatic status for games in this respect is simply to make explicit what authors like John Searle, John Rawls and others did when they relied so heavily and interestingly upon the structure of games to clarify their ideas on institutional facts.

One need only read Rawl's article to see that he was indeed using games as a (at least one) paradigmatic instance of a practice. But on the other hand he believes the practice conception to be especially relevant in understanding "legal and legal-like arguments" (1955: 43 n. 27). Hence, we could start by trying to ascertain how naturally Rawl's conception of practice would fit the law on the one hand, and games on the other.

"Practices are set up for various reasons, but one of them is that in many areas of conduct each person's deciding what to do on utilitarian grounds case by case leads to confusion." *This* reason for setting up a practice seems clearly applicable to the law. Instead of having people autonomously deciding how much income tax to pay, it seems more convenient authoritatively to lay down the rate of income tax. Games are typically set up for other reasons. In the case of games, the practice is created *ex novo* for the sake of having one more activity to engage in.[1] Hence, here we have two reasons (there might be more: "various reasons" seems to suggest more than two) why practices are set up. So far so good.

Since "practices are set up for various reasons", those reasons mentioned by Rawls do not characterise the notion of practice. It is probably the fact that they are set up that is characteristic of institutions. For the time being I will bypass this issue, to which I will return shortly. One consequence of the creation of a practice is "the specification of a new form of activity; and from this one sees that a practice necessarily involves the abdication of full liberty to act on utilitarian and prudential grounds". This seems straightforwardly applicable to games and the law (though a complication will shortly appear). In both cases the agent is not supposed to act on the basis of an all-things-considered judgment, rather she is supposed to follow the rules. This must be understood as meaning that, from the point of view of the practice, it defines what is to be done in particular cases. Both citizens and players might find themselves in situations in which reasons external to the practice might require them to break the rules of practices. Normally, however, and unless the practice itself recognises an exception for those cases, the presence of those reasons will not constitute an excuse from the point of view of the practice. This is obviously true of games: think of a football player asking the referee to validate a hand-goal of his because he promised his dying son to score a goal (*see* Detmold, 1984: 49). In order to explain the idea of the agent's "abdication of full liberty to act on utilitarian and prudential grounds" in practices, Rawls uses the following example:

> In a game of baseball if a batter were to ask, "Can I have four strikes?" it would be assumed that he was asking what the rule was; and if, when told what the rule was, he were to say that he meant that on this occasion he thought it would be best on the whole for him to have four strikes rather than three, this would be most kindly taken as a joke. One might contend that baseball would be a better game if four strikes were allowed instead of three; but one cannot picture the rules as guides to what is best on the whole in particular cases, and question their applicability to particular cases as particular cases" (1955: 38).

[1] Bert Roermund showed me that I had to be more careful here: of course, we do have further reasons to invent this new activities. For example, games are usually contests, they typically involve some kind of competition between players. However, they are not reasons for a particular game, but reasons for a game of such-and-such features. There is a subtle distinction to be drawn here: "the value or point of chess [i.e. winning] is certainly artificial; chess constructs by its rules its unique method of winning. But the value of winning in general is an inseparable part of play [contest] and no more artificial than play itself is" (Detmold, 1984: 160–1).

Is this feature of games shared by the law? I believe it is not. "It would be better if I were allowed four strikes, hence I am allowed four strikes" shows without doubt that the speaker has not mastered baseball, or else that he is joking. In the context of a baseball game, that utterance is simply nonsense. Here it is indeed the case that "if one wants to do an action which a certain practice specifies then there is no way to do it except to follow the rules which define it" (Rawls, 1955: 38). But "it would be better if a legatee were not entitled to the legacy if he's been convicted for the murder of the testator, hence he is not so entitled" is radically different. Whether or not this is a good argument will depend, of course, on the peculiarities of the legal practice in question, but one cannot say that, in general, the mere fact of putting it forward shows ignorance or lack of seriousness about the law.

Here we should remember the context of Rawls's article. In it, the distinction between two concepts of rules was designed to defend utilitarianism "against those objections which have traditionally been made against it in connection with punishment and the obligation to keep promises" (Rawls, 1955: 21). Rawls's strategy was to claim that there were some spheres of social life that were, so to speak, insulated from the *direct* application of moral considerations. But by focusing upon those paradigmatic instances (games), he overemphasised that feature of practices, and tended to regard complete insulation as defining the notion of practice. When rules are practice-rules, "a player in a game cannot properly appeal to [moral] considerations as reasons for his making one move rather than another" (Rawls, 1955: 31). Indeed, "it is *essential to the notion of a practice* that the rules are publicly known and understood as definitive", because "those engaged in a practice recognize the rules as defining it" (*ibid.* 36, emphasis added).

So the picture we get is as follows: first, the notion of a practice is defined in such a way that it obviously includes, if anything at all, developed legal systems: "I use the word "practice" throughout as a sort of technical term meaning any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defences, and so on, and which gives the activity its structure" (Rawls, 1955: 20 n.1). Secondly, the notion of practice is developed as showing that it *essentially* implies the insulation of some sphere of activity from the application of moral considerations (in this case, since the article is a defence of utilitarianism, the utilitarian principle); thirdly, and precisely because of the perfect isolation that characterises them, games are placed at the centre of the analysis, as "paradigmatic" instances; and fourthly, the conclusions thus obtained are said to belong to the *notion* of a practice, with the consequence that they clearly apply to the law, since the law obviously constitutes an instance of a practice in the sense defined at the beginning. But the problem is that the perfect insulation that characterises games is due to the *particular sort of practice* they are, and thus *that* aspect of games is not essential to the notion of a practice. It is certainly not shared by the law.

Let me explain this last point, to make it clear that I am not begging any question (I will come back to this point, hence all I need do here is to make this claim plausible enough for the reader to suspend disbelief and go on). A game-like insularity would imply a highly formalistic approach to law. I do not want to claim, of course, that such an approach is necessarily wrong (indeed we shall see that we can find instances of that formalist approach, and that that backs my general argument). But in the case of games such a formalistic approach is something *about* the practice that any would-be participant has to understand *before* actually engaging in it, while in law the correctness of such an approach is a substantive claim to be argued and defended inside the practice, as one more ordinary, first-order legal claim. To dispute the complete insulation that characterises games is to show lack of understanding about the practice of them; but to dispute that very same approach to the interpretation and application of rules of law is not to display ignorance regarding the fundamentals of legal practice, but to defend a substantive legal claim and in a particular case to advance moral arguments to ground an interpretation of the law is not to display ignorance or lack of seriousness about that practice.

In other words, one could say that a category is missing in Rawls's distinction. Rawls does say that his distinction is "not intended to be exhaustive" (1955: 40). Notice, if correct, what the argument so far implies and what it does not imply. Rawls's attempt to develop a distinction between a "summary" and a "practice" conception of rules is defective insofar as the notion of practice is supposed to be especially "relevant to understanding legal and legal-like arguments" (Rawls, 1955: 43 n. 27). But as a defence of utilitarianism, it might (or might not) well be the case that both kinds of practices (i.e. games and the law) are sufficiently insulated from moral considerations to allow for the distinction between "justifying a practice and justifying a particular action falling under it" (*ibid.* 20), which is Rawls's main purpose. If we are to follow Rawls's advice, however, and apply the distinction to the understanding of law and quasi-legal argument, we shall need a more sophisticated distinction, one that (ideally) is able to explain the difference in the insularity displayed by different practices. This is what I shall try to develop in this chapter.

Before that, however, it will be useful to refer to another aspect of Rawls's argument. What is the correct understanding of rules, the summary or the practice view? According to Rawls, that is the wrong question to ask. The point is not that either the practice or the summary conception has to be true of all rules: "Some rules will fit one conception, some rules the other; and so there are rules of practices (rules in the strict sense) and maxims and "rules of thumb"" (Rawls, 1955: 40).

But the problem is, how are we to know whether a particular rule is to be interpreted according to the "summary" or the "practice" view? Rawls believes that there might be cases in which "it will be difficult, if not impossible, to

decide which conception of rules is applicable. One expects borderline cases with any concept, and they are specially likely in connections with such involved concepts as those of a practice, institution, game, rule and so on" (Rawls, 1955: 40). But can we make sense of the idea that it is sometimes *impossible* to decide which conception is applicable? "Summary" and "practice" rules are applied in radically different ways; what participants must know and master before being able to engage in each kind of practice is significantly different. In the "summary" conception, a rule is a rule of thumb: it has no normative force of its own (here a rule is *quae rem quae est breuiter enarrat* (*Digest* 51.17.1), something which briefly *describes* how a thing is: *see* below, at 151ff). In this conception an agent can reasonably apply a rule only if she has reason to believe that the proper application of the standards of which the rule is a summary leads to the same solution offered by the rule.[2] In the practice conception the agent is not supposed to act on her assessment of the situation beyond the assessment needed to establish that the rule applies. It follows that no agent can really participate in a practice unless she is able to decide whether the rule is a summary or a practice rule. An observer trying to understand what the agent is doing, on the other hand, might find it impossible to determine whether the agent treats the rule as a summary rule or as a practice rule. Indeed, we are in such a position concerning at least significant aspects of Roman law (for the full argument and examples, *see* below, at 141ff). The importance of this point for a theory of legal reasoning can hardly be overemphasised.

THE GAME-ANALOGY

An idea underlying Rawls's general notion of "practice" is, as we have seen, that both games and the law are correctly regarded as practices in his sense. Now, for Rawls this might not be a problematic point, since he was not concerned with providing an analysis of law, but that idea is also recurrent in contemporary legal theory. Among many other authors, as we shall see, H L A Hart seems to have shared this view and, in *The Concept of Law*, relied heavily on the similarities between them. In fact, he relied upon that analogy *so* heavily that it is not too big an exaggeration to say, with Judith Shklar, that games were "Hart's obsession" (Shklar, 1986: 105), or that "H L A Hart described law as a complex game" (Morawetz, 1992: 16).[3]

[2] Notice that this is not necessarily because the agent has actually applied those standards to the case and decided that the solution is the rule's solution. It might very well be that, e.g., the agent does not have time to consider how those standards apply, and so she relies on the previous decisions summarised by the rule. Still, the rule acquires its force from its being a good summary ("the law may not derive from a rule, but a rule must arise from the law as it is" says the *Digest* 51.17.1), and a conscientious agent would always try to check the accuracy of that summary, circumstances permitting (of course, in many cases circumstances will not).

[3] On Hart's "obsession", in addition to Hart, 1994: 310 (index entry for "Games"), see Hart (1953: *passim*).

One of the most important functions of the game-analogy in *The Concept of Law* was to help Hart in ascertaining the rights and wrongs of formalism and rule-scepticism. One of the rule-sceptic's arguments, Hart tells us, is based upon the fact that

> [a] supreme tribunal has the last word in saying what the law is and, when it has said it, the statement that the court was "wrong" has no consequences within the system: no one's rights or duties are thereby altered . . . . Consideration of these facts makes it seem pedantic to distinguish, in the case of a supreme tribunal's decisions, between their finality and their infallibility. This leads to another form of the denial that courts in deciding are ever bound by rules: "The law (or the constitution) is what the court says it is" (1994: 141).

To answer this argument, Hart considered what a game would be like if one were to see games as the sceptics see the law. According to him, such a game would not be like any ordinary game, but a rather odd one he called "scorer's discretion". In it, "rules" are mere predictions of what the referee will do, since they are what the scorer says they are. To see the law as "scorer's discretion", however, is a mistake, Hart claimed, for the same reason a normal game like football or cricket is *not* "scorer's discretion". Though it is strictly possible that any game may be transformed into "scorer's discretion", this possibility does not imply that all games are, actually, "scorer's discretion": "[t]he fact that isolated or exceptional official aberrations are tolerated does not mean that the game of cricket or baseball is no longer being played" (Hart, 1994: 144–5).

I do not want to discuss the whole of Hart's argument against the rule-sceptic (which might be right regardless of the flaws in the "scorer's discretion" argument), but only to note that this particular argument is not very convincing. The sceptic could answer by saying that in games people do not disagree about what the rules are, nor about how should they be applied. They might discuss whether or not Maradona used his hand to score his famous goal against England (though nobody would still like to deny that), rather than whether a particular player's touching the ball with his hands "counts" as a hand ball: is it not an amazing fact that however passionate participants and spectators can be (as is all too well-known nowadays, in Europe, particularly in England) no serious disagreement exists as to what the rules mean and what they demand?

It seems, though, that Hart would not agree with this. He believed that rules of games were as open-textured as any other rule. In games the scoring rule, "though it has, *like other rules*, its area of open texture where the scorer has to exercise a choice, yet has a core of settled meaning" (1994: 144; emphasis added).

It is at least arguable that this assimilation of rules of games to legal rules, on which Hart relied so heavily in his discussion of formalism and rule-scepticism in chapter VII of *The Concept of Law*, distorts the way in which rules of

games are applied: they are not controversial (except in special cases, like the "dangerous play" rule, as we shall see). If this is correct, then there is no reason why the rule-sceptic must be committed to a denial of the difference between football and "scorer's discretion", and Hart's argument (at least his analogy with games) would become harmless. If, as Paul Valery has said, "no scepticism is possible where the rules of a game are concerned" (quoted in Huizinga, 1970: 30), then the analogy cannot be used against *legal* rule-scepticism, where scepticism is (to say the least) possible.

In other words, Hart thought that whatever was true concerning rules of games was also true concerning legal rules by virtue of the fact that in both games and the law "rules" are an important element. There is no obvious reason why the sceptic has to go along with this unstated premiss and Hart did not offer a non-obvious one.

As was said at the beginning of this section, however, Hart was not alone in thinking the game-analogy to be useful for the analysis of legal concepts. Ronald Dworkin's case is interesting for two reasons: on the one hand, he uses the game-analogy at two crucial moments: first, when introducing his (now famous) distinction between rules and principles, and then, to present his (also famous) thesis of what he called "the interpretive attitude". On the other hand, while in the first case he was interested in the *similarities* between games and the law, in the second his point was to *distinguish* one from the other. And in both cases the feature of games he relied upon was the same, i.e. the certainty of their rules.

In *Taking Rights Seriously*, he argued that there was a logical distinction between rules and principles, because only the former were applicable in an all-or-nothing fashion. Though this all-or-nothing aspect of *legal* rules might not be obvious, it "is seen most plainly if we look at the way rules operate, not in law, but in some enterprises they dominate—a game, for example" (Dworkin, 1977: 24).

The interesting point, in my view, is that the reason "this all-or-nothing feature is seen most plainly" in the case of games *is the very same reason* why "scorer's discretion" is so different from cricket or football: because of the certainty of the application of the rules of games. The game-analogy was meant to throw light on something important *about the law*, but it was (in both cases) based upon a feature of games that the law *does not share* (namely, its complete insulation from moral considerations). Here again, the unstated assumption is that "legal rules" are, so to speak, the same kind of entity as rules of games. But Dworkin himself later *distinguished* the law from games when he wanted to explain what is to have an "interpretive attitude". He argued that the two components[4] of the interpretive attitude were independent from each other, so that participants could accept one without necessarily having to accept the other: "We do that in the case of games and

[4] The two components are "the assumption that the practice . . . has some point", and that "the requirements of [the practice] are sensitive to its point" (Dworkin, 1986: 47).

contests. We appeal to the point of these practices in arguing about how their rules should be changed, but not (except in very limited cases) about what their rules now are; that is fixed by history and convention" (1986: 48).[5]

According to this passage, Dworkin would probably say that precisely because no interpretive attitude is adopted either by players or spectators of, say, football, the interpretation and application of the rules of football can have the very high level of certainty they do have. This, however, creates the problem of establishing in which way what we call "rules" in games are the same sort of thing that we call "rules" in law. Dworkin claimed that "legal rules" were all-or-nothing standards, and that this "is seen most plainly" if we look not to legal but to game-rules. There was, for Dworkin in 1967, a common feature between rules in games and rules in the law: they were both *rules*, i.e. standards that were "applicable in an all-or-nothing fashion".

Twenty years later, however, we were told that there was, after all, a fundamental difference between law and games: only the former is an "interpretive concept". So we can legitimately wonder, does this fundamental difference affect the "all-or-nothingness" of legal rules? Maybe the rules of games are all-or-nothing standards not because of a feature they have in virtue of their being rules, but because players and spectators have developed no interpretive attitude towards football. If this is true, then the conclusion would be that in Dworkin's definition there is no such a thing as a legal rule (more on this later, at 98).

The fact is, we are told both by Hart and Dworkin, that some (quite important) features of rules are easier to see if we look at games, but harder if we are looking at the law. This, however, does not necessarily mean that those features are to be equally found in the latter with only an extra effort of observation; maybe they can be easily seen in the former only because they cannot be seen at all in the latter.

INSULATION

Consider the case of the (now not-so-) recent modification of the offside rule in football. As is known, "a player is in an offside position if he is nearer to his opponents' goal line than both the ball and the second last opponent": "a player is not in an offside position if he is levelled with the second last opponent or with the last two opponents". A player in any of these last two situations would have been in an offside position under the old offside rule. We are told that after the 1990 World Cup played in Italy, in which most teams adopted highly conservative and defensive strategies (therefore diminishing the

[5] The "very limited cases" Dworkin had in mind in this passage are probably cases like the one he discussed in *Taking Rights Seriously*. These cases are not counter-examples to my argument, as I will try to show below (for the argument, *see* below, at 31ff; for Dworkin's very special case, *see* below, at Ch. 2, n. 19.

quality of a football match as a spectacle), FIFA modified the rule in an effort to make the game more aggressive and more attractive to audiences and so forth.

The following I take to be rather obvious: in the period between (some moment before the end of) the 1990 World Cup and the issuing of the new offside rule, all of the reasons for having the new rule existed, but they were irrelevant at the moment of applying the old rule. They are equally irrelevant *after* FIFA's decision: an umpire could not, after the new rule had been introduced, decide that he should apply the rule in the light of its goal, and hence that he was going to be (say) "less strict" in the application of the rule in those cases in which the involved play was likely to be part of an aggressive strategy, even if we are prepared to imagine that in so doing the referee would be likely to increase the desired effect of the new regulations. After the decision was taken, the reasons for it were as irrelevant in the context of adjudication as they were before. The very suggestion that the referee could apply the rule in this way seems to be nonsensical[6] (in fact, it is not clear what "less strict" could mean in this context).[7] Thus the rule is completely insulated from the reasons for it.

But in legal adjudication things are different. This will be discussed in considerably more detail below, but for the time being suffice it to say that lawyers do speak of interpretations being more or less strict, and the idea that a law should be interpreted in the light of its purpose is all too common. Needless to say, this is not the *only* kind of argument that can be used to interpret a law (some will say that it is not even a good argument), but for the time being, I only need to claim that this *makes sense*, in a way that the offside argument does not. The idea that interpretation should be purposive is not necessarily controversial insofar as it is limited: nobody would deny that legal rules should *sometimes* be interpreted in the light of the goals they are supposed to advance. What is controversial is whether or not this is *always* the case. But we need not adopt this strong Fullerian position to see that a football referee is not, except when the rules explicitly grant him that power, supposed to consider the purpose of the rules at the moment of applying them. This general feature of adjudication in games is absent in the law: even according to a positivistic theory of law there will be cases in which "assumptions about the purposes the rule is meant to advance would take a prominent—perhaps even pre-eminent—role in solving the particular difficulties encountered" (Marmor, 1994: 154).

The point is that in games the application of a rule is always straightforward, while in law, at least sometimes, the application of an otherwise valid and clear

[6] In this context, to say of an argument that it is nonsensical is to say that the fact of a speaker seriously offering that argument would be taken by others either as a joke or as proof that the speaker is not really playing, or does not understand the game, etc. This was Rawls's point discussed above (at 3).

[7] I am ignoring some complexities of the offside rule, like the so-called "passive" offside. They are not an objection to the thesis presented here, any more than the existence of "discretion-granting" rules like the dangerous play or advantage rules are (*see* below, at 31f).

rule can be contested. Notice that this claim does not amount to saying (not *now*, anyway) that there are no cases in which the application of legal rules is not as clear as the application of rules of games. My claim here amounts to the rather obvious observation that, while it would not lift a lawyer's eyebrow to notice that the application of a clear general rule to a particular case can sometimes be problematic, to find that referees and players disagree about what a penalty, or a goal, or a hand ball are, would indeed surprise any football fan.

In other words in football (as in all games) the fact that a rule does not make some aspect x of a concrete case relevant means that x is irrelevant for the application of that rule.[8] Part of what you need to learn when studying football is, for example, that during the course of the game the ball can only be handled by the goalkeepers in their penalty boxes, and that under no circumstances can a different player or the goalkeepers outside that area intentionally use their hands to move the ball. There might be disagreement about whether or not a particular touching of the ball was intentional, but a player that was to offer moral reasons to justify his handling of the ball would show, as Rawls claimed, that he does not understand, or else is joking.

The latter example was consciously selected for the discretion that the rule allows the referee to determine whether a particular hand ball was "intentional" or not. A rule against "touching" or "handling" the ball might be vague in the sense that it does not clarify whether an unintentional handling of the ball would count as a hand ball (is it a case of a player touching the ball or of the ball touching the hand of a player?). Since whether or not the semantic meaning of rule requires intentionality is not clear, there will be cases in which people will disagree about whether a player's touching the ball counts as a hand ball.[9] Insofar as rules of a game rely on the normal use of words, there will always be space for this kind of vagueness. Thus, *this* kind of vagueness is common to games and the law.

[8] Consider the following case: on Saturday, 13 February, 1999, Arsenal played Sheffield United at Highbury. The score was 1-1 with just over 10 minutes of the game to play. A Sheffield United player, Lee Morris, went down after a challenge by an Arsenal attacker. Since the referee did not stop the game, a Sheffield United player kicked the ball out. Morris recovered, and the match restarted. Arsenal's Ray Parlour tried to throw the ball to United keeper Alan Kelly to give the visitors unchallenged possession, but his team-mate, Nwankwo Kanu, hunted the ball down and slid a low cross into the path of Arsenal's Marc Overmars who then scored. Sheffield United failed to equalise and the match finished 2–1 for Arsenal. What is important here is that the goal was valid, though it was scored in violation of one of the most clear and undisputed requirements of *fair play*. The fact that Arsenal's goal was grossly unfair was irrelevant for its validity, because the scoring rule in football does not make any reference to the fairness of the scoring. The match was later ordered to be replayed, which is another way of emphasising the same: the possibility of simply "invalidating" the goal because of Arsenal's unfairness was not discussed (I am grateful to Kevin Walton for bringing this case to my attention).

[9] Remember the penalty that led to Italy's equaliser against Chile in their 1998 World Cup match (probably only we Chileans remember it, so here it is: seven minutes before the end of the match a Chilean defender, Ronald Fuentes, touched the ball with his hand inside Chile's penalty box (or the ball hit Fuentes' hand) and the referee awarded Italy a penalty kick. The decision was controversial, since the intentionality of Fuentes' hand ball was in question. Italy scored and the match ended 2–2). For the record, the official FIFA rule *does* require intentionality (Law 12).

Am I not conceding Hart's point here? Not really. We only need to look at the way in which rules of games are applied in *real matches* to see how restricted the vagueness warranted by this argument really is. The claim that the rules of football and legal rules are *equally* vague is so descriptively inaccurate that one wonders if it should receive any serious consideration. The problems that arise when applying *legal* rules to concrete cases go well beyond problems of uncertainty at the (semantic) borderline, and this is almost as old as the law itself. I will offer only three cases taken from different periods in a two thousand-year span:

(1) Roman jurists knew that not to take Rawls's batter's attitude at least sometimes was legally mistaken. In other words, for them not to take into account considerations other than those needed to ascertain whether or not the rule's operative facts were fulfilled would have demonstrated lack of mastery of the law, i.e. exactly the opposite of Rawls's batter. Consider just one example, a piece written by Paul, who lived in the third century AD:

> D.1.3.29 (Paul, *libri singulari ad legem Cinciam*). *Contra legem facit, qui id facit quod lex prohibet, in fraudem uero, qui saluis uerbis legis sententiam eius circumuenit* (it is a contravention of the law if someone does what the law forbids, but fraudulently, in that he sticks to the words of the law but evades its sense).

In other words, it is not possible to know whether by following a rule we are following the law unless we can ascertain the *ratio* (sensus) *legis*. Ascertaining the *ratio legis* supposes the ability to consider how some moral considerations bear on the issue. Marc Overmars did not infringe the rules of football when he scored his goal against Sheffield United, since the scoring rule in football was silent concerning *fair play* (see above at n. 8); but it is at least arguable that the court would be misapplying the law if it grants the legacy to the murderous legatee on the basis that the statute of wills is silent on the legatee's killing of the testator (see below, at 34ff).

(2) The second example dates back to 1688. Discussing the interpretation of laws, Samuel Pufendorf comments upon a Bolognese case:

> there was a law of Bologna, that whoever drew blood from another person in a public place should suffer the most severe penalties. On the basis of this law a barber was once informed upon, who had opened a man's vein in the square. And the fellow was in no little peril because it was added in the statute that the words should be taken exactly and without any interpretation.

This example is offered by Pufendorf as an illustration of his claim that

> When words, if taken in their plain and simple meaning, will produce an absurd or even no effect, some exception must be made from their generally accepted sense, that they may not lead to nothingness or absurdity (1688: Book V, Ch. 12, § 8, pp. 802–3 [547]).

Here we can find Pufendorf facing a legal problem and doing as a matter of course what for players and referees, when facing a similar problem in a football game, is always wrong: i.e. arguing that, since there is some value *beyond* the rule for which the rule is an instrument, (how could it be known that a given result is "absurd" if not by reference to such an external value?) the question of how the rule is to be applied to particular cases has to be answered taking that into account. And to take that into account (some) moral considerations have to be referred to. In this case Pufendorf is indeed behaving like Rawls's batter; he is claiming that since it would be better on the whole if the rule had an exception for barbers, the rule does have an exception. And one might think that Pufendorf is wrong, but if so, he is wrong as a matter of Bolognese law: we are not entitled to conclude that he did not understand the law or that he was joking.[10]

(3) Much the same can be said about many of the cases imagined by Lon Fuller in the second half of the twentieth century. I will only refer to one of them here. We are invited to consider the existence of a rule to the effect that 'It shall be a misdemeanour, punishable by fine of five dollars, to sleep in any railway station", and to imagine that

> two men are brought before me (i.e. the judge) for violating this statute. The first is a passenger who was waiting at 3 am for a delayed train. When he was arrested he was sitting upright in an orderly fashion, but was heard by the arresting officer to be gently snoring. The second is a man who had brought a blanket and pillow to the station and had obviously settled himself down for the night. He was arrested, however, before he had a chance to go to sleep (1958: 664).

Notice that in this and in Pufendorf's case there is no doubt as to the meaning of the words used by the rule. Anyone who was to say that the barber should not be punished or the businessman should not be fined because the former didn't "really" draw blood in a public space or the latter wasn't "really" sleeping in the station, would demonstrate lack of mastery of English. In fact, the problem is created rather than solved by the fact that the case is indeed uncontroversially covered by the semantic meaning of the rules.

This is important because there has been a strong tendency in contemporary legal theory to regard the issue of the defeasibility of legal rules as one that has to be tackled in terms of the general defeasibility of concepts.[11] But any sensible explanation of legal reasoning will find a form of defeasibility that is not reducible to semantic defeasibility. The cases discussed by Pufendorf and Fuller, and the danger we are warned against by Paul are not cases of what we

[10] This last sentence should not be interpreted as relying on Pufendorf as an authority concerning the law. The argument is not "look, Pufendorf was so clever, he could not have been joking or misunderstood the law". It is, rather, that Pufendorf is doing something that no lawyer would regard as off-limits.

[11] This is indeed what Hart himself, only sometimes I believe, thought: *see* Hart, 1994: 124ff; the issue is discussed below (at 89ff).

might call "semantic" defeasibility. In the case of Ronald Fuentes' hand ball we might disagree because we disagree about the case being covered by a rule against "intentionally" handling the ball, and we might not be sure about Fuentes' intentionality. But in the cases discussed by Pufendorf and Fuller this is not the reason why we might not be sure about the application of the rule. The truth is, we are unsure on the face of the fact that the cases can easily be shown to be covered by the semantic meaning of the relevant rules. This is a different sense in which a rule can be defeated. My claim is that *this* form of defeasibility, characteristic of legal reasoning, is not to be found in games. From now on, I will refer to this form of defeasibility every time I use that word without further explanations.

It is clear that the question of which legal cases will cause problems of this kind depends on the peculiarities of the legal practices involved (see Atiyah and Summers, 1987, as discussed below, at 207ff) . What in my view cannot be denied, however, is that the law as we understand it is a kind of practice in the context of which this is a (more or less) common problem. One could even say that mastery of the law (what law students are supposed to learn before becoming lawyers) *is* (or at least includes) the ability to recognise these cases, while mastery of football is (or at least includes) to understand that the application of a clear rule cannot be discussed during a match.

If this is correct so far, then it would naturally follow that an explanation of legal reasoning, or a criticism (like Hart's) of rule-scepticism cannot be based on an analogy between games and the law. A good starting point for an explanation of legal disagreement is, therefore, to give a closer to look at the similarities and differences between the two.

A GENERAL THEORY OF INSTITUTIONAL FACTS[12]

Both games and the law figure profusely in the literature on institutions. Both games and the law are (or allow for) paradigmatic instances of "institutional facts". I think that there is an important truth here, but that truth is obscured when some crucial differences (we have seen one; we shall see more shortly below) between the two are disregarded. The argument offered in the last section was designed to give this point initial plausibility: in some important sense rules seem to be much more well-behaved, so to speak, in games than in the law. If this feature of rules in games is not taken account of, the game-analogy can easily backfire: after all, it might be the case that *precisely* because rules of games are certain in their application when rules of law are not, that the rule-scepticism Hart was arguing against is right, or that *precisely* because

[12] This section benefited from Professor John Searle's detailed comments and criticism for which I am grateful (every now and then in the text I refer to what Searle said or did not say "at the Buffalo conference", meaning the Marvin Faber Conference in Applied Legal Ontology held at Buffalo, NY in May 1998).

of this certainty in the application of the rules of games, *legal* rules are not all-or-nothing, as Dworkin claimed.[13]

I want to claim that games and law are institutions of different *kinds*. A convenient way of developing the argument would be, therefore, to begin with what we could call a "unified" theory of institutional facts like goals, contracts, hand balls and the like. This is now possible since John Searle has recently offered what he called "a general theory of institutional facts" (1995). I will try to show that a theory that does not recognise the existence of two *kinds* of institutions cannot but fail to account for some peculiarities of the neglected kind.

Regulative and Constitutive Rules

In an often-quoted passage, Searle introduces his now famous distinction:

> I want to clarify a distinction between two different sorts of rules, which I shall call *regulative* and *constitutive* rules . . .. As a start, we might say that regulative rules regulate antecedently or independently existing forms of behaviour; for example, many rules of etiquette regulate inter-personal relationships which exist independently of the rules. But constitutive rules do not merely regulate, they create or define new forms of behaviour. The rules of football or chess, for example, do not merely regulate playing football or chess, but as it were they create the very possibility of playing such games (1969: 33).

Searle believes that this reflects an "intuitively obvious distinction" between two different kinds of rules. He himself acknowledged that he was "fairly confident about the distinction, but do[es] not find it easy to clarify" (*ibid.*). However obvious that distinction looked to Searle, it proved controversial. Among others, Anthony Giddens[14] has argued that "that there is something suspect in this distinction, as referring to two types of rule, is indicated by the etymological clumsiness of the term "regulative rule". After all, the word "regulative" already implies "rule": its dictionary definition is "controlled by rules'" (Giddens, 1984: 20). In other words, all rules can, in one way or another, be said to be regulative. This is, naturally, no objection to Searle's distinction, since he does not claim that constitutive rules *do not* regulate (notice, in the following displayed quotation, his qualification of "*purely* regulative" rules, and also his claim, in the previous quotation, that "constitutive rules do not

[13] I do not want to pursue these arguments here, since they are not important for the point discussed in the main text. It might be the case that, in the end, Hart is right against the rule-sceptic or Dworkin is right in his claim about the logical distinction between *legal* principles and rules. In both cases, however, once account is taken of the uncontroversial nature of the application of the rules of games, the game-analogy ceases to be a supporting reason (i.e. as it was used supporting Hart's or Dworkin's arguments) and provides the reader with a (not necessarily conclusive) reason to believe exactly the opposite.

[14] I begin with Giddens' criticism because he is the only critic to whom reference is made in *The Construction of Social Reality* (at 230 n. 10).

merely regulate"). "Constitutive rules"—this is Searle speaking at the Buffalo conference— "of course regulate behaviour, but they do something more, they create the possibility of forms of behaviour that would not exist without those rules."

This should be readily granted. But if all constitutive rules do regulate behaviour, then it cannot also be the case that *all* rules are also constitutive, since in that case there would be no distinction whatsoever. Are all rules, then constitutive? Searle answers:

> There is a trivial sense in which the creation of any rule creates the possibility of new forms of behaviour, namely, behaviour done in accordance with the rule. That is not the sense in which my remark is intended. What I mean can perhaps be best put in the formal mode. Where the rule is purely regulative, behaviour which is in accordance with the rule could be given the same description or specification . . . whether or not the rule existed, provided the description or specification makes no explicit reference to the rule. But where the rule (or system of rules) is constitutive, behaviour which is in accordance with the rule can receive specifications or descriptions which it could not receive if the rule or rules did not exist (1969: 35).

Some critics of Searle have not been convinced by this argument, and have pressed the point that all rules are, really, *both* constitutive *and* regulative. One of them is Joseph Raz, who invites us to compare the following two pairs of act-descriptions:

1 (a) "Giving £50 to Mr Jones" (b) "Paying income tax"
2 (a) "Saying 'I promise'" (b) "Promising".

In Raz's view,

> descriptions 1 (a) and 2 (a) specify acts which are in accordance with the rules in a way which could be given regardless of whether or not there is such a rule. Therefore, the rules are regulative. Descriptions 1 (b) and 2 (b) describe actions in accordance with the rule in a way that could not be given if there were no such rules. Therefore, the rules are constitutive, as well. Since for every rule one can formulate a similar pair of act descriptions, all rules are both constitutive and regulative (1992: 109).

But here there is a clear *non sequitur*. From the fact that in both 1 and in 2 "one can formulate a similar pair of descriptions" it does not follow that that is the case "for every rule". This is the more obvious when we notice that Searle would probably not object to Raz's claim that both taxes and promises are institutional facts. Searle would (rightly) claim that one can think of (other) rules for which the (b) item of the pair is missing. In other words, it is not the case that all actions in accordance with rules, because of that very fact, admit of this dual description. That *is* the case concerning tax law and promising, but not concerning rules of, say, the decalogue. The following is not a complete pair of act-descriptions:

3 (a) Honouring one's father and mother (b) (empty)

since, though in the relevant circumstances "giving £ 50 to Mr Jones" and "saying 'I promise'" counts as paying taxes and promising, 'honouring one's parents' does not count as anything. Hence not *all* rules are constitutive.

What about

4 (a) Being nice towards one's parents (b) Honouring one's parents?

In this case the description contained in 4 (b) is a form of *appraisal* rather than a specification (Searle, 1969: 36; *see* also Cherry, 1973: 302). If instead we had something like

5 (a) Not honouring one's parents (b) To be guilty of a sin

then the rule in question (the Fourth Commandment, Exod. 20:12) would indeed, I believe, be constitutive. But this would depend upon the (contingent, i.e. not necessarily implicated by the Fourth Commandment) existence of the institutional concept of *sin*.

Notice that the concept of sin is not conceptually (as opposed to theologically, as the case might or might not be) needed in order either to understand or to apply the Ten Commandments: they can be understood as simply stating what it is right and wrong to do. The institutional concept of sin is born, so to speak, when someone offers an interpretation of a (up-to-then-not-institutional) practice in terms of institutional facts (see below at 25ff). Notice further that, if we introduce the institutional concept of sin, not only the fourth, but *all* of the Ten Commandments suddenly become constitutive: each of the acts described by each commandment (taking the name of the Lord in vain, killing, committing adultery, stealing, etc.) becomes an X term to which the institutional Y term "sin" is attached.

Perhaps this is the gist of Raz's critique. Maybe he should be understood as saying that concerning legal rules in developed legal systems, there will be always a description available for the (b) item. This is, I believe, true, but it fails to follow that all rules are both regulative and constitutive. What follows is that a rule acquires its character (regulative or constitutive) from the normative system to which it belongs.

We can now go one step further. Since all rules are regulative, but some are also constitutive, it follows that some rules are purely regulative. But this, I believe, presents an interesting question: can "purely regulative" rules exist in the context of institutional systems ("institutions", following Searle for a while, being "systems of constitutive rules")? Are there purely regulative (say) legal rules, for example?

Tony Honoré, for instance, has claimed that a satisfactory theory of individuation of laws must allow for the following kinds of laws:

1. *Existence laws* create, destroy or provide for the existence or non-existence of entities.
2. *Rules of inference* provide how facts may or must or should preferably be proved and what inferences may or must or should preferably be drawn from evidence.
3. *Categorising laws* explain how to translate actions, events, and other facts into the appropriate categories.
4. *Rules of scope* fix the scope of other rules.
5. *Position-specifying rules* set out the legal position of persons or things in terms of rights, liabilities, status and the like.
6. *Directly normative rules* (which are few in number, but important) guide the conduct of the citizen as such (Honoré, 1977: 112; on the importance of a theory of individuation of the law, which gives the background of Honoré's claim, *see* Raz, 1980: Ch. 4).

I hope it is clear that items 1 to 5 cannot be purely regulative. Rules of the third type, for example, "A young person is any person who has attained the age of 14 years and is under the age of 17 years" (Honoré, 1977: 102) are plain instances of constitutive rules, i.e. rules of the form "X counts as Y in C" (Searle, 1995: 43ff). The same can be said of items 1, 4 and 5. Rules of inference, I believe, are also typically constitutive: they specify what counts as evidence for the existence of an institutional fact like a contract or a will.

This leaves only 6, "directly normative rules". Are they not purely regulative? They certainly do not *constitute* what they regulate: if they did, they would fall into another category. The problem is, these rules cannot be purely regulative since they are expressed in institutional terms. They are almost tautological (Searle seems to believe that they are, in fact, tautologies: 1969: 191).[15] As Honoré argued,

> the fact that criminal legislation by and large defined what constitutes an offence and does not directly forbid the obnoxious conduct . . . reveals . . . that the directly normative rules of a modern system are for the most part platitudinous generalities. "Do not commit an offence". "Abstain from torts". "Perform contract". "Pay debts". "Discharge liabilities", "Fulfil obligations" . . .. These basic norms are not tied to specific act-situations, and this confirms, if it needed confirmation, how unsatisfactory would be any general programme of individuating laws on the basis of act-situations. But of course the norms presuppose for their application in legal discourse that the system contains rules which do specify the act-situations falling within the general categories "offence", "tort", "contract", "debt", "liability", "obligation" (Honoré, 1977: 118).

[15] Though I will not pursue this matter further, it is interesting to notice that it is dubious whether they can be tautologies. Barry Smith argues, following Adolf Reinach, that in (what Searle calls) "systems of constitutive rules" one must "eventually arrive at basic institutional concepts [BICs], which is to say: institutional concepts not capable of being further defined on the institutional level" (Smith, 1993: 318). They are not capable of being defined in non-circular ways in terms of non-institutional concepts, since then "all institutional concepts would turn out to be thus definable". This reinforces my conclusion that there is no space for purely regulative rules in "systems of constitutive rules". *See* Sergot *et al.* (1986), for an attempt to translate an actual piece of legislation, the British Nationality Act, into a set of definitions (i.e. rules of the form "X counts as Y in context C"). One could then take "citizenship" to be a basic institutional concept (or to be definable in BICs, or to be definable on the basis of concepts that are in turn definable on the basis of BICs, etc.). The nature of these BICs raises problems I need not pursue: Smith goes on to say that the only explanation available to Searle would be to accept that truths about BICs "express irreducible material necessities of the Reinachian sort, that is, express necessary relations between certain uninventable *sui generis* categories" (1993: 318–9; Smith's reference is to Reinach, 1913).

Given that the act-situations these rules regulate are constituted by other rules, i.e. those defining "tort", "contract" and the like, these rules cannot be purely regulative. A rule like "perform contracts", for example, seems to be regulative, but it imposes on a party to a contract a "negative power" (which corresponds, naturally, to the other party's "positive power" of requiring compliance). This negative power (and its correlative positive) is necessarily part of any specification of the Y term in any rule of the form "X counts as Y in context C" when "Y" stands for "contract". Hence it is not only (part of) a constitutive rule, it is a rule without which nothing we would recognise as a contract could exist. This highlights an important feature of Searle's distinction: the distinction is not one between rules, but one between *systems of rules*.[16]

Searle does not agree. He believes that it is perfectly possible *both* for constitutive rules to exist without belonging to any system of rules (a kind of "stand-alone" constitutive rule) *and* that there can exist purely regulative rules in the context of institutions (we shall see how these two points are, really, the same). But I fail to see how he can allow for these possibilities without giving up the distinction altogether. At the Buffalo conference Searle offered the following examples:

(1) *Stand-alone constitutive rules.* "You might have a tribe that has a procedure for selecting a leader: he who can lift the biggest stone, like the Vikings. The one who can throw the rock the farthest, he is the boss. So there you got one constitutive rule. And there is no whole system. It's just that they recognise him as the boss".

The problem is, "being the boss" might receive a non-institutional interpretation. "Being the boss" can be a brute fact if it means that I recognise someone as the boss because she has shown that she is the strongest, and because of that it is better for me to do as she wants until I get as strong as she is (here "boss" works as a mere label for "strongest"). This involves no institutional fact (*see* Searle's remark on labels, quoted below), and hence "she's the boss" does not make reference to any constitutive rule.

Things are different, however, if by "she's the boss" I mean or imply that in

[16] To be sure, he recognises the fact that, usually, "constitutive rules come in systems" (1969: 36; *see* also 35). But constitutive rules are constitutive *because* they belong to institutions, not when they are floating in some kind of normative vacuum (*see* Cherry, 1973).

some sense I have the *duty* to do as she commands. *Now* we're talking institutional. By being chosen as the boss, she has acquired a new status that allows her to demand obedience from the rest of us: X counts as Y in context C. This normative language is made possible by constitutive rules. How can we distinguish the two situations? Well, we have to look at precisely what "being the boss" means. If "boss" is an institutional fact, we would expect to find not only a rule about how to choose the boss, but also rules that specify the consequences of being the boss: rules that confer powers to the boss and obligations to the subordinates, to say the least. If we have only a rule saying how we can find out who is he or she whose wants we better satisfy or else, then I do not see how we could say that "being he or she whose wants we better satisfy or else" can be an institutional fact. Indeed, it looks to be a good example of one of Rawls's "summary" rule. Contrariwise, if Searle wants to grant that "being he or she whose wants we better satisfy or else" is an institutional fact, then it seems likely that with some ingenuity we shall be able to make any regulative rule into a constitutive one, and the distinction would collapse.

Searle, I believe, saw this problem (the problem of all rules becoming trivially constitutive in this way) in *The Construction of Social Reality*, where he said that

> As I am using the formula (i.e. the formula "X counts as Y in context C") it would not be a statement of a constitutive rule to say "objects that are designed and used to be sat on by one person count as chairs", because satisfying the X term is already sufficient for satisfying the Y term, just from the definition of the word "chair". *The "rule" does not add anything but a label* so it is not a constitutive rule (Searle, 1995: 44, emphasis added).

In the situation imagined by Searle, in which we have only one rule specifying that he or she who can lift the biggest stone shall be the boss, there is no constitutive rule, but a mere label ("boss" a mere label for "strongest"), unless we find further rules specifying the status that goes with the Y term. It follows that precisely insofar as constitutive rules stand alone, they cease to be constitutive rules.

This is not just an amusing detail about institutional facts. The key difference between a label and a status is that only the latter is characterised by a set of powers or functions that are attached to it; hence *it is necessarily the case* that, insofar as there is only a rule saying "he or she who can throw the rock the farthest is the boss", without any function or status being attached to a person's being the boss, and *precisely because there is only one such rule*, "being the boss" would not count as an instance of an institutional fact. Indeed, Searle appears to grant this when he talks about money:

> But to describe these bits of paper with the Y term "money" does more than provide a shorthand for the features of the X term; it describes a new status, and that status, viz. Money, has a set of functions attached to it, e.g. medium of exchange, store of value, etc. (1995: 46).

(2) *Purely regulative, institutional rules.* At the Buffalo conference Searle offered rule 48 of Baseball as an example of this: "no player, coach or manager shall dispute a decision of an umpire". But this rule is not purely regulative, since it contributes to defining the positions of players, referees and coaches in the game of baseball. A person's having the status of referee in baseball is an institutional fact that is defined (among other things) by the fact that no player or coach can dispute his decisions as to how the rules are to be applied. If rule 48 is not constitutive of what a referee (a player, a coach) is, then it is difficult to see how powers can be ascribed to institutional statuses.

We see how points (1) and (2) are linked: a constitutive rule of the form "X counts as Y in context C" must indicate not only how something or someone gets to occupy the Y position (i.e. not only how you go about choosing the boss, hence the Viking rule either was not constitutive or did not stand alone), but it also "has to assign a new *status* that the object does not already have just in virtue of satisfying the X term" (Searle, 1995: 44). This status gets assigned by further rules specifying what the powers of the status-holder are. Among those rules, in the case of baseball, we shall find rule 48. Rule 48 cannot be a purely regulative rule.

Before pursuing this point any further, let us go back to the beginning and consider the criteria offered by Searle to distinguish regulative from constitutive rules. In *Speech Acts* Searle offered two different criteria: one at p. 33 and another at p. 35 (I will call them "33" and "35", respectively. Both of them were quoted above, at 15f). What is the relation between these two criteria? I think it can be shown that they do not necessarily coincide, because they answer different questions. As Geoffrey Warnock said:

> This supposed distinction between "two sorts" of rules is really, I think, a confused groping after two other distinctions. There is, first, a distinction between two ways of saying what people do—one way which, as for instance walking, or hitting balls about, or waving flags, involves no reference to any rules, and another which, as for instance playing tennis, or signalling, or bequeathing property, does essentially make reference to rules, or presupposes them. Then, second, there is a broad and rather woolly distinction between two different "objects" of rules, or reasons for having them. It is not the *object*, presumably, of the criminal law to "create the possibility" of committing criminal offences, though it incidentally does so; the object is to "regulate" in certain respects the conduct of members of society. By contrast, while the rules of, say, soccer do "regulate" the way in which balls are kicked about in fields, it is in this case the *object* of (some of) the rules to "constitute" a certain exercise in physical skill and ingenuity, to "create" a particular game for people to play (1971: 38).

Now having Warnock's idea in mind, let us consider the two criteria of Searle's distinction. As should be remembered, the first criterion (33) was that while regulative rules regulate "antecedently or independently existing forms of behaviour", constitutive rules create or define *new* forms of behaviour, like playing chess or football.

Strictly speaking, as Searle saw, what *creates* the possibility of playing chess or football is not *a* rule, but a *system* of rules ("the *rules* of chess..."). We have already seen that no single rule can allow for institutional facts. Hence this is not a criterion to distinguish constitutive from regulative *rules* but *systems of rules*. Some systems of rules exist, as Warnock said, in order to "regulate in certain respects the conduct of members of society", i.e. to regulate antecedently or independently existing behaviour. The point of some other systems of rules is, on the other hand, to "create" particular activities (both Warnock, as we have seen, and Searle (1995: 50) agree that while chess is an instance of the latter, criminal law is one of the former).

The second criterion (35) looked at the description of behaviour which is in accordance with the rule. When the rule is constitutive, the action in accordance with it can be given a description which would not be available if the rule did not exist; concerning regulative rules, this is not the case.

Now, this criterion tells a rather different story: criminal law, for example, is typically regulative in the first sense, a point that, as we saw, is uncontroversial. Criminal law does not exist *in order to* create the possibility of committing offences, but to regulate antecedently existing forms of behaviour. But in (not necessarily too) developed legal systems, the rules of criminal law are necessary to describe, for example, that Jones is "guilty" of "murder in the first degree" though he is "excused" by "mitigating circumstances" etc. Hence, according to the second criterion, the rules of criminal law are constitutive (*see* MacCormick, 1998: 335, where he argues that "the boundary between regulative and constitutive is unclear in Searle's schema"; *see* also MacCormick and Weinberger, 1986: 23, where they claim that "particularly unsatisfactory is the Searlean distinction between constitutive and regulative rules").

Searle's new general theory of institutional facts is still liable to this problem. This time the distinction makes its appearance in the book with the help of the following pair of examples: "drive on the right-hand side of the road" (regulative) and "the rules of chess" (constitutive). Here we can see Searle using the first criterion. The "drive on the right" rule is said to be regulative because it regulates driving and driving is an antecedently existing form of behaviour (Searle, 1995: 27), while rules of chess are constitutive because they "create the very possibility" of playing chess (notice again the singular of the former as opposed to the plural of the latter).

Now consider for a moment Searle's new paradigmatic regulative rule: "drive on the right-hand side of the road". If the rule's literal formulation is (something like) "drive on the right-hand side of the road, or you shall be forced to pay £5" the rule does not create "new possibilities of behaviour" and is, therefore, (purely) regulative. But if the rule's formulation were "failure to drive on the right-hand side of the road shall constitute an offence" (as it is, in fact, likely to be), it would indeed be creating such a new possibility (to wit, to commit an offence), and it would be constitutive.

I do not believe Searle would like to accept that the selfsame rule can be regulative or constitutive according to its literal formulation. But the only way in which this can be avoided is to focus not upon the rule, but upon the whole system of rules to which the rule belongs: if it is a legal rule in a developed legal system ("developed" here excludes "systems of primary rules" in the sense of Hart, 1994: 91), then it will be constitutive, since it will single out one form of behaviour as the X term to attach to an institutional Y term like "being guilty of an offence", with a particular set of negative and positive powers. If it is only a rule of etiquette (like the "stand on the right" rule that applies in the escalators of the London Tube), then it would be purely regulative (but it could even be a rule of a weird game: "driving on the right counts as scoring one point").

As mentioned before, Searle does not agree with my talk of two criteria. He claimed (at the Buffalo Conference) that 33 is a definition (rather than a criterion), and 35 simply a "pedagogical device", another way of looking at the issue, but this does not seem to be more than a verbal disagreement. Searle does agree that one consequence of the existence of institutional facts is that one can use something one might want to call "institutional language". I would also like to distinguish 35 from 33. I would like to say that 33 is the important idea, and that whether or not one would use institutional language to describe institutional facts (i.e. whether or not 35 obtains) depends, to an important extent, upon technical details about the canonical formulation of the rules involved. But whether or not a system of rules is institutional in 33's sense is not something that 33 makes dependent upon the canonical formulations of the involved rules. As we have seen 33 is not a definition that singles out a constitutive rule, but a *system of rules*. The distinction is, therefore, between systems of rules that have as their main point the creation of a new activity and those whose main point is to regulate a pre-existing practice. Systems of the first kind, however, do regulate pre-existing forms of behaviour (e.g. the rules of football regulate the ways in which players can get the ball moving), and systems of the second kind do constitute new forms of behaviour (e.g. to be guilty of an offence). The distinction contained in 33 is not based on the fact that some systems constitute and others regulate, but on the fact that some systems regulate pre-existing forms of behaviour in order to create a new activity, while others create the possibility of new forms of behaviour in order to regulate some pre-existing form of behaviour.

This is not a particularly strong objection to Searle's original claims: for some purposes it might be of use to focus upon particular rules only. Indeed, since both systems do constitute (though in different directions: one constitutes in order to regulate and the other regulates in order to constitute), if all we want to talk about is the fact that institutional language introduces a special ontology, it might be enough simply to talk about "systems that allow for institutional facts" and in this sense we might legitimately refer to both games and law, to institutions that regulate-to-constitute and those that

constitute-to-regulate. The argument contained in this Chapter is not designed to show that Searle's original distinction was mistaken. It is, rather, that the possibilities that are opened up for the analysis of the "institutional ontology" are missed if one stops simply at the point where one can glance at institutional facts. My claim, to speak metaphorically, is that one can use more sophisticated glasses, glasses that allow one to see the differences between the inhabitants of this institutional world. Leaving the metaphor behind, the claim is that to put systems of rules, rather than rules, under the spotlight, provides far greater insight into the way rules work.

In the next section I will try to provide some examples of this last claim. I will try to show that, because he uses a distinction that is designed only to show whether institutional facts are possible, when it comes to explaining the features of that institutional ontology, Searle tends to ascribe to all institutional fact features that in truth belong to some of them not because they are institutional, but because they are the particular sort of institution they actually are. Thus, he is led to distort one kind of institution by forcing upon it the features of another.

A Critique of Searle's General Theory

The Evolution of Institutions

The first problem I want to discuss is related to the issue of the evolution of institutions. Can an institution evolve without the participants being aware that they are evolving one?

Searle's answer is, indeed they can. Consider the example of money. People can go around buying, selling and exchanging, without their thinking that the particular goods they use as a medium of exchange is "money":

> The evolution may be such that the participants think, e.g. "I can exchange this for gold', "this is valuable", or even simply "this is money". They need not think "we are collectively imposing a value on something that we do not regard as valuable because of its purely physical features", even though that is exactly what they are doing . . .. In the course of consciously buying, selling, exchanging, etc., they may simply evolve institutional facts (Searle, 1995: 47).

Now, why is it possible for people to evolve institutional facts without being aware of it? The answer is that they can keep doing what they were doing all along, and the institution will grow, so to speak, on the back of the practice. As Zenon Bankowski has argued, concerning promises:

> the institution comes about because gradually a practice grows up where, for example, we do something we say we will, not merely because of the substantive reasons we had in saying we would do it, but also because of the reason that we said we would do it. At first that is one among all the reasons but gradually it excludes the others and so we might say the convention of promising grows up. We do it because we promised and the other reasons are excluded. Thus the institution grows up on the back of the substantive reasons since the reason that it is a promise can be seen as the universalisation of the substantive reasons (1993: 13; *see*, for a similar point, Atiyah, 1981: 120).

Before proceeding any further, notice that Bankowski is *not* trying to offer a reductionist analysis of promising in terms of whatever substantive reasons the parties might think they have for promising. There is no need to emphasise that any such reductionist account cannot be a complete analysis of promising. I think that Bankowski's point here should be understood as aiming at the same target as Searle's assertion (1995: 50f), that "in many cases the X term is chosen precisely because it is supposed to have the features necessary to perform the function specified by the Y term", though "even in these cases, something is added by the Y term". Bankowski's claim, from this point of view, is that the Y term "grows up on the back" of the substantive reasons *normally* behind the X term, and that it adds to it *some degree* of insulation from the actual presence or absence of those substantive reasons in a particular instance of the X term. Under normal circumstances, we have good substantive reasons to grant an agreement (the X term) the binding force of a promise (the Y term). Therefore, we treat agreements as binding promises without having to check, in every instance, whether those substantive reasons are actually present.

This growing of the Y term on the back of the X term, however, is something that can only happen regarding institutions that "constitute to regulate", that is, institutions that create the possibility of institutional facts because of the improved regulatory effects this *technique* allows. Because the institutions of criminal law are not necessary to sustain the practice of punishing people for failing to behave according to what Hart called "primary" rules (in much the same way in which we saw that the concept of sin was not conceptually necessary either to understand or to apply the Ten Commandments), those who administer the punishments need not think of the rules of criminal law in constitutive terms (in an "undeveloped" system, it could be enough to have a list of "do's" and "don'ts"; or, rather, a list of "don'ts—or else"). They can simply continue the practice of punishing people, and at some point in time a writer (what in Scotland, for example, is called an *institutional* writer)[17] can offer an interpretation of the practice of punishing

[17] See Cairns (1994: 90): "In France and Spain, institutional works were obviously linked to attempts to create and to promote a unified national law. This cannot be so for Scotland, since Scots law was unified. It is, however, worth considering that Scots law did require unification in a different sense, in that the separate constituent parts of the law—customary, Canon, Roman and statute—had to be worked into a convincing whole; and this unification of the law into a general Scots law is generally taken to have been carried out by the Scottish institutional writers, especially Stair. The disparate elements of Scots law are connected with the various different jurisdictions—royal, heritable and ecclesiastical—and it must be of importance in this respect that in the seventeenth and eighteenth centuries all the various jurisdictions tended to be united into one centrally organised system of justice."

people in terms of institutional facts (*see* MacCormick, 1974: 62f; 1998: 333; MacCormick and Weinberger, 1986: 12).[18] But when the institution is one that "regulates to constitute" (one that specifies how things are to be done in order to create a new activity e.g. how a ball is to be kicked about in fields in order to create the game of football) it cannot evolve on the back of the practice, since without the institution there is no practice at all. There cannot be a pre-institutional practice of football, in the sense in which it is possible for a pre-institutional legal practice to exist; a Hartian "regime of primary rules" of football is, I believe, a conceptual impossibility (*see* Amselek, 1988: 209: "it is impossible to imagine that one can play a game without implying that one is following the corresponding collections of rules") . The first group of people who thought of football, for example, must have been aware of the fact that they were imposing a particular meaning on three wooden posts that did not have that meaning by virtue of their physical characteristics.[19]

(When I was a boy we used to play football in a park. As there were, of course, no goal posts in the park, we had to use our jumpers and bags as goalposts. The first time my friends started to throw their bags and jumpers around I could not understand what were they up to, until one of them said: "this is your goal, and that is ours": everything was clear from then on. We could not have played football in the park had we not been aware of the fact that by placing those bags and jumpers where we placed them we were collectively assigning meaning to them, a meaning that was not exhausted by the physical properties of the bags and jumpers. But the POWs who, in German concentration camps, as the standard story goes, started to give and accept packages of cigarettes in exchange for other goods need not have been aware of the fact that by their giving and taking cigarettes in those circumstances they were assigning to cigarettes a meaning not exhausted by their physical characteristics).[20]

Systematic relationships between institutional facts

One feature of institutional facts, according to Searle, is that they "cannot exist in isolation but only in a set of systematic relationships to other facts" (1995: 35). For money to exist, a system of exchange has to exist beforehand, and for a system of exchange there has to be a system of property and

[18] I am not saying that the writer *creates* institutional facts where there were none; she makes explicit what was up to then implicit in the practice: this is Searle's point. What we have is a seamless process from pure brute facts to implicit institutional facts to explicit institutional facts (for an illustration of this process in legal history, *see* Cairns, 1994; Stein, 1983) .

[19] But they could, couldn't they, think that the posts had some magical feature, so that football was something that had to be played in those terms because of broader considerations (such as the aim of not insulting the Gods, etc.)? This answer is not available to Searle, who would not be willing to call this "game" a game (1995: 36n): "to the extent that professional sports have such [broader] consequences, they cease to be just games and become something else, e.g. big business".

[20] See Wonnacott and Wonnacott (1990: 38–41) for the text-book version of this story. Wonnacott and Wonnacott follow Radford (1945).

property ownership. "Similarly, in order that society should have marriages, they must have some form of contractual relationships. But in order that they can have contractual relationships, they must have such things as promises and obligations" (*ibid.*)

Generally speaking, the existence of systems that "constitute to regulate" presupposes the existence of the practice the system is created to regulate. This is, obviously, because the point of the development of the system is its regulatory impact on the practice. But this shows that this (i.e. the fact that some institutions presuppose other institutional and non-institutional facts) is not the case by virtue of some mysterious characteristic of institutional facts, but because of the particular kind of institutional facts under consideration. Concerning games, again, the point is less straightforward.

Searle, however, thinks that games are not counterexamples to his claim, though "it might seem" that they are, "because, of course, games are designed to be forms of activity that do not connect with the rest of our lives in a way that institutional facts characteristically do". When this point is looked at carefully, Searle claims,

> even in the case of games there are systematic dependencies on other forms of institutional facts. The position of the pitcher, the catcher, and the batter, for example, all involve rights and responsibilities; and their positions and actions or inactions are unintelligible without an understanding of these rights and responsibilities; but these notions are in turn unintelligible without the general notions of rights and responsibilities (1995: 36).

It is not clear whether Searle thinks that *baseball* is unintelligible without such notions as rights and responsibilities or, as he later claims, that this fact (the fact that baseball so depends) is a consequence of games generally "employ[ing] an apparatus—of rights, obligations, responsibilities, etc.—that is intelligible only given all sorts of other social facts" (*ibid.*, 56). In any case, this does not seem to be the case. We can understand, make sense of, and even *play* chess without knowing a thing about the "apparatus" used in India during or before the sixth century (or wherever and whenever it was invented: that we do not need to be sure of its origins to play is another way of making the point). Baseball and football are played all over the world, and that is not a proof that the notions of "rights, obligations and responsibilities" are common to the human race at large, unless one wants to hold on to Searle's point and claim that the fact that we can understand games is a proof of a shared "apparatus" between human beings of all times and places (a weird argument, would it not be, for a natural law doctrine?)

Granted, today we would use some idea of responsibility to understand the different functions of, say, a goalkeeper, a defender, etc. in a football team, but without such notions one is still able to play football. Those notions seem to me to be linked more to the idea of a successful strategy than to the very notion of what football is.

Finding in games this feature, which is obviously present in legal concepts (money, marriage and the like) forces Searle to weaken his requirement. At the Buffalo Conference, Searle's explanation for our ability to understand chess in the face of our ignorance about sixth century India was that "we do have to know that they spoke a language, and that they understood such things as, the queen has more power than the pawn . . .. The basic institutional form is language; the basic institutional move is the speech act, and that's universal".

This is, of course, something I do not want to deny. But notice how weak in content this requirement is when compared to the heavily loaded requirement that "in order to . . . have money, that society must have a system of exchanging goods and services for money. But in order that it can have a system of exchange, it must have a system of property and property ownership" (Searle, 1995: 35). We are not talking about language here, we are talking of social structures, production relations and the like: in order to have institutions, some social structures must be in place. When it comes to games, all we are told is that the society must have a language. This is common, says Searle, to all institutional facts, and I do not want to deny it. Using Barry Smith's language (Smith, 1993), we might say that all institutions stand in a relation of ontological necessity to language; but some of them (like those Searle referred to in the first paragraph of section 4 at page 35 of *The Construction of Social Reality*) also stand in such relations to much more concrete and contingent practices (like private property and the like). This is something that is relevant for the ontology of different kinds of institutional facts, something that is missed by Searle's too rough conceptual apparatus.

Institutions and their Consequences

So let us go back to Searle's statement in *The Construction of Social Reality*, where he claims that "It might seem that games are counterexamples to this general principle, because, of course, games are designated to be forms of activity that do not connect with the rest of our lives in a way that institutional facts characteristically do" (1995: 36). To the best of my knowledge he does not explicitly refer to this characteristic of institutional facts elsewhere in the book, and it is not clear what he has in mind. One characteristic way in which institutions (i.e. systems of constitutive rules) connect with our lives is that they allow us to do things that we could not otherwise do: we can promise, we can play football and so on. But in this sense games *do* connect with our lives in the same way, hence this is not the sense in which Searle intends his remark on page 36 (where he claimed that games *do not* connect in the way institutional facts characteristically do). The sentence that immediately follows the one discussed here seems to imply that the way in which institutional facts characteristically connect to our lives is that the former have *consequences* for the latter: "Today's philosophy department softball game need have no consequences for tomorrow, in a way that today's wars . . . are intended precisely to have consequences for tomorrow".

Hence, the fact that institutional facts have "consequences for tomorrow" is characteristic of them. This is why the caveat above concerning the paradigmatic *status* of games as institutional facts is important: we have seen that, concerning two important features of institutional facts, games are unlike the institutions that are really important. Indeed, I do not want to object to the thesis that *some* institutions stand in systematic relationships with other institutional and non-institutional facts. In fact, this is an extraordinarily important feature of institutions like the law. But this is not a characteristic of institutions *qua* institutions (since there are institutions that have no systematic relationships to other facts), but only of institutions that "constitute to regulate", *because* they do so. The reason for this is simple: since the institutional (i.e. constitutive) apparatus is used to regulate a practice that exists independently, that apparatus must, of necessity, have "systematic relationships" with the institutional and non-institutional facts that are part of the practice to be regulated.

I take Searle's point of institutions "having consequences for tomorrow" to be his way of singling out what I have been calling institutions that "constitute to regulate" from those that "regulate to constitute". My last claim can, therefore, be expressed in Searle's terms by saying that institutions have "systematic relationships to other facts" *because* they have "broader consequences": games do not have consequences, hence they need not have those relationships. Indeed, insofar as games do develop those relationships, Searle himself believes that they "cease to be just games" (*cf.* 1995: 36).

Constitutive and Regulative Institutions

I agree with Searle when he says that the important criterion to characterise institutions is the first one (i.e. 33). According to it, the law is a "regulative" institution, since its point is to regulate antecedently existing forms of behaviour (and to do that in a better and more efficient way it creates the possibility of new forms of behaviour). Games, on the other hand, are "constitutive" institutions, that is, systems of rules whose point is to create new possibilities of behaviour rather than to regulate antecedently existing forms of it (though they doubtless do regulate some pre-existing forms of behaviour in order to do this). A distinction of this kind is obviously behind Ronald Dworkin's claim that

> chess is, in this sense, an autonomous institution; I mean that it is understood, among its participants, that no one may claim an institutional right by direct appeal to general morality . . .. But legislation is only partly autonomous in this sense (1977: 101).

Thus, it transpires that the important distinction is not based on whether the *rules* are constitutive of institutional facts or regulative of pre-existing forms

of behaviour, since we might find rules of both kinds in either system. Thus "thou shalt not drive on the left" would be regulative (because it is not needed to describe the action), but "it shall be an offence to drive on the left" counts as constitutive (since it *is* required to describe the action of committing an offence).

Because "regulative" (regulatory) institutions are justified by their regulatory effects (i.e. by their "broader consequences") those effects have an impact upon the application of the rules. In regulative (regulatory) institutions the rules set out only what is "presumptively" the case, and the fact that that presumption can be defeated in concrete cases allows for problems like those discussed by Paul, Fuller and Pufendorf. In "constitutive" (autonomous) institutions, since the institution is not justified by its regulatory effects, consideration of those effects need not affect the application of the rules, which can (but need not) be indefeasible: here we go back to the initial observation, i.e. the fact that disagreement about what is the law is a common phenomenon while it is most *un*common concerning games. The defeasibility of *legal* reasoning, then, is a consequence of the *kind* of institution the law is understood to be (we shall soon see that some forms of ancient law can be said to have been "autonomous": *see* below at 49ff). But to see this, to understand legal reasoning, we need a theory of institutional facts that can account for this distinction.[21]

A note about the word "institution". As should by now be evident, I am using this word in a loose sense. Or rather, I am using it as defined by Searle, as "systems of constitutive rules" (1969: 51). I understand "constitutive rules" in this definition as meaning "rules that provide for the existence of institutional facts". Therefore, both "constitutive" (autonomous) and "regulative" (regulatory) institutions are in this sense institutional: both of them allow for the existence of institutional facts (an example of a non-institutional system of rules is Hart's "régime of primary rules" in Hart, 1994: 91ff). My reason for using the word "institution" in this sense is to emphasise the fact that what I take to be the true distinction between the regulative and the constitutive is not the fact that only the former regulates and only the latter allows for institutional facts. Both kinds of institutions do both, but in different directions, so

[21] After the next paragraph I will cease to talk of "constitutive" and "regulative" institutions. I believe that the argument presented in this section is best viewed as a way of taking Searle beyond Searle, that is to say, of building upon Searle's "general theory of institutional facts". But even if my argument fails as a critique of Searle's views, I still think it has intrinsic value. For this reason, from now on I will label "autonomous" institutions those that "regulate to constitute", like games (i.e. those systems of rules that if my argument is correct correspond to Searle's "constitutive rules"). The other kind (i.e. those that "constitute to regulate") I will call "regulatory" institutions; they would correspond to Searle's regulative rules. In choosing these labels I have tried to give them a Searlean flavour, while at the same time suggesting that they represent a different (i.e. hopefully improved) version of Searle's two kinds of rules. Beyond that there is nothing to be read in the labels. They could be replaced by "A-" and "B-institutions" (in fact, labels of this latter kind were used in a previous draft of this chapter, and I am grateful to Professor David Garland who suggested to me the labels I am using now).

to speak: one regulates in order to constitute, the other constitutes in order to regulate.[22]

However, there still remain some objections that could be presented against the thesis that there is a difference in *kind* between two models of institution. These objections take the form of alternative explanations for the differences between legal and game-adjudication, explanations that would not be committed to the claim that they are qualitatively different. To them we should turn now.

THE GAME OF LAW

In this section I want to address some objections to the thesis presented above, objections that amount to the claim that the difference between institutions like games and institutions like the law is not one of kind, but one of (at most) degree. Needless to say, since the argument up to now has effectively claimed precisely the contrary, I have to show why all these objections fail.

To begin with, however, it could be said (i) that I have overvalued the certainty of norms of games. Is it not the case that some norms of games are, after all, indeterminate in a Hartian sense? Any football fan knows that some actions are *core* instances of, say, *dangerous play*, but also that the referee will have to exercise discretion to decide whether or not some actions—which can be said to be penumbra instances of "dangerous play"—are to be punished (my stipulation concerning the word "defeasibility" above at 14 goes some way towards answering this objection). Furthermore, (ii) the fact that these controversial applications do not generate the same controversy as hard cases in law might be due to the existence of a secondary rule of adjudication in football according to which decisions must be produced on the spot and without further consideration (indeed, it is very difficult to imagine a game like football without such a rule).

In my view, however, both of these facts are explained, not by the reason that natural languages are necessarily open-textured, but by the existence of rules to that effect. With regard to (i), the use of vague standards like "dangerous play"[23] is, himself tells us, a particular legislative "technique"

[22] Again, this is basically a stipulative definition, and for that reason it is important to see its implications. It follows Searle's stipulation, but it would not be agreed upon by, e.g. Neil MacCormick, who claims that a definition of institution in terms of constitutive rules "would simply involve an obvious confusion between the law of contract and the legal institution "contract" itself which is regulated by that branch of the law" (1974: 51). It would also commit one to say that a contract is a different institution in Germany than in France, while it could at least be claimed that the (same) institution of contract exists both in German and in French law, though subject to different rules. I believe that MacCormick is right in making the distinction between the system of rules and the institution that exists under it, but for ease of exposition I will use one word to refer to both, hoping that the context will make the precise meaning clear.

[23] "An indirect free kick is awarded to the opposing team if a player, in the opinion of the referee . . . plays in a dangerous manner" (Law 12).

(1994: 132) that it is reasonable to use when "it is impossible to identify a class of specific actions to be uniformly done or forborne and to make them the subject of a simple rule" (*ibid.*). This must be distinguished from the philosophical claim about language according to which "whichever device . . . is used to communicate standards of behaviour, these . . . *will* prove indeterminate" (*ibid.* 128, emphasis added), since, even if the latter claim were false, that "legislative technique" could still be useful in many cases (thus, we shall see that the Hartian open texture thesis can receive two wildly different interpretations). With regard to (ii), it clearly cannot be the case that we have no disagreement about what the rules of football are for concrete cases *because* the referee has the final say on the matter, since if that were the case we would not be playing (or watching or talking about) football but some form of "scorer's discretion".

But maybe a more sophisticated version of this argument could be advanced along lines suggested by Neil MacCormick. He first noticed and then tried to offer an explanation for what he called the "variable practical force" of rules (1998: 316–7). He argues that rules are of *absolute application* if the "O[perative] F[acts] must be attended unfailingly by N[ormative] C[onsequence], and *NC* may not be put into effect except when either *OF* obtains or some other rule independently providing for *NC* is satisfied by virtue of the ascertained presence of its operative facts"; of *strict application* if "the person charged with applying the rule and managing the activity within which the rule has application is given some degree of guided discretion to make exceptions, or to override the rule, in special, or very special cases"; and of *discretionary application* "if the decision-maker is expected to consider every case in the light of all factors that appear relevant".

Now, what, according to MacCormick, determines the kind of rule a rule belongs to? "The answer is obvious—it depends not on the content of the first-tier rules about a practice, but on second-tier norms laying down the terms of authorization or empowerment of the decision maker" (*ibid.* 317).

The variable practical force of rules is an important feature of them (and fatal to any account of rules as exclusionary reasons, as we will see), but if my argument is correct, MacCormick's explanation cannot be enough. The "second-tier" rules arise when a person is appointed to monitor the application of the rules of the practice (MacCormick, 1998: 312). But MacCormick's own example of rules of absolute application (rules of chess) shows that the practical force of rules is determined even in the absence of second-tier rules. Indeed, Hart (1994) noticed that "many competitive games are played without an official scorer: notwithstanding their competing interest, the players succeed tolerably well in applying the scoring rule to particular cases; they usually agree in their judgements, and unresolved disputes may be few" (at 142). It is the nature of the institution, what determines the practical force of the rules of it (needless to say, second-tier rules, when they exist) that can affect the practical force of rules. But even when they purport to do so, the

nature of the institution sets them an important limit, since they also belong to it, *see* below at 118ff.[24]

So let us look elsewhere for an explanation. It could be argued, to support the thesis that the difference between games and the law is not one of *kind*, but rather one of *degree*, that this difference of degree is explained by the difference in the *complexity* of the regulations that games and the law involve: games restrict the reality they deal with, and so they create a world artificially simple (Huizinga, 1970: 28). The law, on the other hand, at least potentially regulates any situation. In a more restricted world, it is possible to predict and to anticipate any problem the application of a rule will present in the future, while this is impossible in law. The law has what, following Emilios Christodoulidis, we could call a "complexity deficit" (Christodoulidis, 1999).

This view appears promising at first sight, but it is wrong, and to see why, compare the two following cases:

Edson's Case. During a football match, each team is allowed to replace a given number of players only (three in the last World Cup). If one team has already made those replacements, it cannot make any further under any circumstance whatsoever. Now, suppose that this is the case, and that one of the players of team *A* (call him Edson) is an extraordinarily good player: the performance of *A* is largely improved when Edson is playing. Now, team *B*'s manager knows this, so he decides to instruct Harald, one of his players, to severely injure Edson. The manager knows that if Harald succeeds he is likely

[24] The distinction between first- and second-tier rules, however, can be of use in a different sense. Every time I have tried to explain the argument contained in this chapter to others, they have felt challenged to try their best to produce counter-examples (memorable discussions about football ensuing). It is an interesting point that the most convincing counter-examples are rules that have as their obvious point to secure the continuity of a match. Bert Roermund offered the best example: a referee has constantly to balance his whistling in accordance with the rules against his responsibility in keeping the match attractive. In deciding whether or not to award a free kick after a very minor fault, the referee might encounter an application problem. My answer to Roermund's example would be that one can distinguish first-tier rules *of* football (like the rule against handling the ball) from second-tier rules *about* football, rules that purport to facilitate the development of the game. The obvious way in which such a distinction could be made would be to say that the first-tier rules define what football is, while the second-tier rules are rules which take football as something already existing and tries to single out some ways of playing as to be preferred. This explanation would fit perfectly well the argument I am developing here, for it would take rules *of* football as constitutive of that new (i.e. non-existing before the rules) activity, football, and rules *about* football as regulating something already existing, i.e. something not created by them. The reason why I confined this discussion to a footnote, however, is that this explanation would commit me to say that what I used to play with my friends at the park was not "football" since we did not consider the offside rule (clearly a rule *of* football rather than one *about* football), and this conclusion seems to me to be rather pedantic. I will not go further into this problem: the distinction might be difficult to pinpoint with complete accuracy, but it seems to me a natural distinction to make. Strictly, a proper counter-example to my claim would be a rule *of* football (or *of* any other game) that were open to challenge by substantive considerations in the absence of a rule of the game granting discretion to the rule-applier (or, as Pufendorf's rule, in the presence of a rule instructing the rule-applier to apply the other rules in a strict manner). I suppose that if FIFA were to pass a rule instructing referees not to balance their whistling against their responsibility to keep the match attractive, then referees would not have any balancing to do. I have not yet heard such a counter-example.

to be sent off, but he also knows that if the injury Edson suffers is bad enough, Edson will not be able to continue playing and, since *A* cannot make further replacements, both teams will continue the game with ten players (with a significant advantage for *B*, since Edson will not be playing for *A*). So Harald breaks Edson's leg in a vicious tackle.

Elmer's Case. Now imagine that Elmer wants his grandfather's money. He knows that his grandfather has made a will in his favour, but he needs the money now (and his grandfather is, alas, very healthy), so he murders him. Imagine further that none of the provisions of the statute of wills said anything about a legatee killing the testator, and all the requirements it does contain for the validity of a will have been fulfilled by Elmer (and Elmer's grandfather). After he has been convicted for the killing, Elmer goes on to claim the inheritance (*Riggs* v. *Palmer*, 115 NY 506, 22 NE 188 [1889]).

Elmer's case is discussed by the jurisprudential literature as a standard example of a hard case. On the other hand, I submit that Edson's cannot but be a clear case, and that any football fan will agree with me if I say that, however reasonable from a moral point of view that might be, the referee *cannot* (without violating the rules of football) allow *A* to make a fourth replacement. I want to argue that *no difference in complexity* can account for this fact (i.e. the fact that the former is or at least can be a hard case while the latter is definitively a clear one). In Edson's case, the rule does not leave any margin to the referee to decide . . . what? to allow Edson's team a fourth replacement? to declare Edson's team the winner?[25] to grant it an extra goal? an extra yellow card for every player in *B*? to send off one additional player of Harald's team, chosen at random? In Elmer's case, however, though the rule does not appear to leave the judge any scope, it does: the possibility of discussing the application of the rule contained in the statute of wills is present, any sensible counsel would see the possible arguments each side could use in court (it does not matter for the time being whether or not these arguments are good enough to carry the day, but only that they are not to be taken as evidence that the speaker does not really understand what lawyers are supposed to do).

To be useful in this context, the recourse to the different level of complexity between a game like football and the law must be related to the (supposed) inability of the law-maker to predict future cases in such a complex normative system as the law is said to be. This was the idea behind Hart's view on the convenience of uncertainty given by our "relative ignorance of fact" and

[25] Bert Roermund advised me to think about this possibility more seriously than I had. The referee could stop the game, in his opinion, and declare Edson's team the winner. But this is because the rules of the game allow the referee to "stop, suspend or terminate the match, at his discretion, for any infringements of the Laws" (Law 5), if he thinks the infringement is sufficiently serious. But it is, of course, not part of my argument that the referee cannot be *granted* discretion explicitly, as in this case. Incidentally, Harald may, of course, be sued by Edson (or sanctioned by FIFA) after the match, that is something for the law, but it is something that, of course, I need not deny.

"relative indeterminacy of aim" (1994: 128). The solution would then be: in a "simple" normative system (e.g. football) the legislator (i.e. FIFA) can predict all possible combinations of relevant facts in the future, so that any participant can safely assume that the solution provided by the rule is the solution actually sought by the authority. Hence, as all participants acknowledge the authority of FIFA, the rule can be applied to any conceivable case without controversy. In a "complex" normative system (e.g. law), on the other hand, participants cannot assume this knowledge on the part of the authority, because, as the reality the law is dealing with is so complex, it is empirically impossible for any legislator actually to predict all of the possible combinations of relevant factual features in the future. So the complexity of the system (strictly, the enormous number of possible combinations of relevant factual features the system purports to take into account) allows space for the following argument: "this rule should not be applied to this case because it was not *meant* to". That would be the reason, on this interpretation, why purposive interpretation is so useful in legal hard cases.

The problem with this approach is simply that there is *no* reason *at all* to assume that a case like Edson's was actually predicted by FIFA (in fact, I would think that Elmer's case is more easily predictable than Edson's). The referee has to do what he has to in Edson's case *not* because he thinks that FIFA so decided (when, at the moment of enacting the replacements rule, it presented to itself the possibility of a case like Edson's), but because he has (given the nature of the game) no other alternative. In other words, the correct solution is correct, not because FIFA wanted this solution for this particular case when it was passing the replacement rule, but because, given some up to now mysterious peculiarity of games as institutions, what (the members of the relevant committee of) FIFA had in mind when the rule was passed is completely irrelevant. This becomes obvious if we notice that even if the referee happens to know that FIFA did not think of this case, his predicament is the same.

Notice how to explain the difference between Edson's and Elmer's cases on the basis of a complexity deficit must necessarily beg the whole issue. In both cases there is a complexity deficit in the sense that for each of them we might feel that there are some features of the case that should be relevant for its correct solution, though they are not picked up as operative facts by the applicable rule. In both cases we might feel that it would be better if the rule were so drafted as to include an explicit exception for the case at hand. In Edson's case, however, the fact that the rule does not contain such an explicit reference is the end of the issue, while this is not necessarily so in Elmer's.

In brief, the fact that reality is infinitely variable has in itself nothing to do with the issue discussed here, because rules are to be applied to those cases that match the operative facts of the rule only: "the legislator does not issue norms for each individual case . . .. His function consists in the creation of general norms, by means of which he resolves generic cases" (Alchourrón and Bulygin, 1971: 30).

In the strict sense, therefore, both Elmer's and Edson's cases are equally solved by the applicable rules. The difference, therefore, is not that the law has a complexity deficit that football does not have: both of them can have it. The question is rather why this complexity deficit is relevant in legal adjudication while it is not in football-adjudication. But this *is* the problem for which we are seeking an explanation.

The argument begs the question even more clearly if we were to say that Elmer's case is more complex because the law does not restrict (as games do) the considerations that can be referred to in adjudication to those explicitly contained in the rules. This is true, but correct though it is, it is not a good answer to our problem here: according to this explanation, the difference would indeed be one of complexity, but then again, the issue would be why cannot the law exclude such considerations? Or rather, why does the fact that a football rule does not mention some feature *X* count as the rule *excluding* that feature, while in the law (at least sometimes) the same fact does not necessarily imply that consequence? Thus, it turns out that the complexity deficit is not really an explanation, but a different way of describing the same problem: *why does the deficit matter in law and not in games?*

In brief, the situation is *not* that because rules of law are not nuanced enough we cannot take them to be final. It is precisely the other way around: because we expect reasonably appropriate solutions from the law, we allow for some "leeway" in the application of legal rules. Thus, Gottlieb puts the cart in front of the horses when he argues that

> What of a model that would eliminate moral judgment of any sort from the judicial role? It is hard to imagine how legal standards could conceivably reach a degree of specificity in all domains that would eliminate every moral dimension of judgment (1994: 16).

It is not because legal standards are imperfect that we need morality to step in as a corrective; it is because we expect legal judgments to display some given moral quality that we read legal standards as allowing for some "leeway". Since we do not understand the rules of games in the same way (because of their different nature as social practices), we do not have the same expectations regarding them. This has nothing to do with the intrinsic quality of the rules (their "degree of specificity"), but with our understanding of the relevant practices.

It could be useful here to consider another possible explanation, one that seems backed by common sense. According to it, the difference between games and the law is that games are not serious or important, so we don't really care about achieving the right result in games. Because we don't really care, we have such a formalist type of adjudication.

There seem to be some important truth in this explanation, but we cannot take it at face value. The reality is, many people would think that what is at stake in (say) some football matches is (for them) more important than many

things that are or may be disputed in court (witness Bill Shankly).[26] I will, however, return to this point below (*infra*, 47f) because there is a sense in which an explanation of this kind can be useful.

A further explanation could be offered on the basis of the *arbitrariness* of some norms: some norms are arbitrary in the sense that the reasons for each of them are not reasons for their content, but only for their existence. They have what Atiyah and Summers (1987: 13) called *content-formality*. On the other hand, most (though by no means all) legal norms have low content-formality: the reasons for having a norm regarding murder are reasons for the content of such a norm as well (that is, to punish murder). Some legal norms are like rules of games in this sense (e.g. some traffic laws), and some rules of games are like legal norms (e.g. the rule of *dangerous play*). And it could be claimed that rules that have a high degree of content-formality cannot but be applied formalistically.

This explanation would explain precisely the point posited at the beginning, i.e. that some norms of games seem to be open-textured in the same way as norms of law. But, conversely, it would seem to imply that the application of (some) legal norms (i.e. those that have high content-formality) is beyond plausible contestation in the same way that the application of rules of games is, and this is the reason why it fails: the interesting feature of rules of games which is in need of explanation is that they are (at least *can be*) indefeasible, while legal norms are always defeasible (though of course undefeated many times). Even a legal norm with the highest content-formality, like the "drive-on-the-left" norm is defeasible.[27]

After Fuller (1958) it is difficult to believe that there are kinds of legal rules that are beyond defeasibility. Fuller taught us that regarding every (legal) rule, cases can be imagined in which doubts would legitimately be felt concerning the application of that rule. And this is the reason why the explanation we are now considering fails: it explains the defeasibility of legal rules on the basis of peculiar features of particular rules (i.e. their level of content-formality), thus implying that some other rules (i.e. those that have high content-formality) are indefeasible.

Let us try one final, alternative explanation. The difference between games and law is not one of kind, but one of degree: on one end of the spectrum one would have highly abstract and formalised games (e.g. chess), then less abstract games like football, then a highly formalistic legal system (like ancient Roman law, as we shall see), then a less formalistic one, and so on.

[26] "Law is more serious than games (the heavy view). What is one to say to that? That it overrates law? Or that it underrates games? Both answers are required" (Detmold, 1984: 160). It has even been claimed that "play" is one of a reduced number of basic human goods (Finnis, 1980: 87).

[27] Consider the problem facing an ambulance driver when he arrives to a traffic jam and realises that the opposite lane is free and nobody is coming that way (or rather the problem of the judge who has to decide if he deserves a sanction for having used that lane).

This is wrong. An intuitive reason why chess might be a better example than football in this regard is due to the fact that football is a game of physical contact, while chess is not. What relevance does this fact have for a theory of institutional facts? Consider the following: a hand ball occurs when a player (other than the goalkeeper in his penalty box) touches the ball with his hands. What is a "hand"? In the rules of football there is no definition of what a hand is, so if a mutant player touches the ball with his fifth limb there might be a problem of application after all. But a bishop in chess has nothing to do with an actual bishop, nor a knight with an actual knight, and so on. The rules of chess completely define what a bishop in chess is. There cannot be a mutant bishop. Then we should expect to find more cases of linguistic defeasibility in football than in chess.

At this point, it may be worthwhile to pay attention to Samuel Pufendorf"s criticism of Hugo Grotius. Here Pufendorf and Grotius were discussing the reason that explains the fact that mathematical knowledge is certain in a way in which moral knowledge is not. Grotius thought that the explanation for this significant difference between mathematics and morality was due to the fact that mathematical concepts are so defined that there is always a crisp and bright line to be drawn between them, while "in moral questions, on the contrary, even trifling circumstances alter the substance, and the forms, which are the subject of enquiry, are wont to have something more closely to this, now to that extreme" (1646: Book II, Ch. 23 § I, p. 557 [393]).

For Pufendorf, however, Grotius's dictum that "in moral questions, even trifling circumstances alter the substance" was ambiguous. In one sense (which Pufendorf labelled "qualitative") this is true, but it is also true of mathematics, "for it is also true that a line which varies in the slightest degree from straightness, tends to curvature", and thus this fact cannot explain the difference between the certainty of mathematical knowledge and the uncertainty of moral knowledge. If Grotius's dictum is to be understood in a different, "quantitative" sense, that is, if

> the saying means that the slightest circumstance increases or lessens the quantity of an action, we answer that this is not always true, at least in a civil court, where the judge often pays no regard to trifles. And even granting this, the fact does not lessen the certainty of moral matters, since even in mathematics the slightest addition or detraction makes a change in the quantity (Pufendorf, 1688: Book I, Ch. 2 § 10, p. 34 [23–4]).

But from this Pufendorf did not conclude that there was no difference between mathematical and moral knowledge. Indeed, he believed that "a certain latitude is found in moral quantities", but his explanation for this was different from Grotius. Pufendorf noticed that "physical quantities can be exactly compared with one another, and measured and divided into distinct parts, because they are in a material way object of our senses", hence we can "determine accurately what relation or proportion they have to one another";

> *But moral qualities arise from imposition,* and the judgement of intelligent and free agents, whose judgement and pleasure is in no way subject to physical measurement and so the quantity which they conceive and determine by their imposition cannot be referred to a like measure, but retains the liberty and laxness of its origin (Pufendorf, 1688: Book I, Ch. 2 § 10, p. 35 [24], emphasis added).

We shall shortly see that the law can attain the high level of certainty games have if it is seen as being part of the "physical" world. Insofar as the law is seen as something that "arises from imposition", as something *regulatory* in character, there is space to question the application of its rules to any particular case, and that introduces "a certain latitude". Further differences in the vocabulary should not bother us at this stage. The point is, when we leave the football-creating convention and start playing football, the rules of the game are seen by players to belong to the structure of the world in the same way in which the "rules" of bridge-building belong to the world for engineers: if you want to build a bridge, do such-and-such; if you want to score a goal, do such-and-such. But consider if you want to write a will, do such-and-such (but actually, if you don't, it might still be the case that you succeed in writing a will; if you do, it might be the case that you fail, etc.).

So one way in which we could express the distinction I have been trying to draw between two types of institution is saying that one kind is supposed to be seen by the participants as "arising from imposition", while the other is supposed to be seen as "simply the way things are". There are different ways in which things might be, and thus we should not be surprised to find out that the permeability of one game to the problem of strictly linguistic defeasibility is greater than the next (there can be mutant football players, but there cannot be mutant bishops in chess). But there is a difference of kind to be made: we shall see (in the last section of this chapter) what are the consequences for legal reasoning that follow from the fact that the law is *not* seen as "arising from imposition".

I have argued that failure to draw a distinction of this kind affects (though of course need not invalidate) John Searle's general theory of institutional facts, even though (maybe precisely *because*) Searle did not deal with the subject of defeasibility. It is about time, therefore, to consider whether (and how) the defeasibility of legal rules is a source of similar difficulties for an "institutional theory *of law*".

LAW AS INSTITUTIONAL FACT

In his inaugural lecture some twenty five years ago, Neil MacCormick put forward the thesis that "if the law exists at all, it exists not on the level of brute creation . . . but rather . . . on the plane of institutional facts". What makes propositions of law true or false, he tells us, is not "merely the occurrence of acts

or events in the world, but also the application of rules to such acts or events" (1974: 51). Contracts, for example, are institutions of law. But legal institutions are not identical with rules, since Chilean contract law is one thing, the contract I have with the University of Talca is another (this point was mentioned above, at n. 22). MacCormick's claim is that institutions are "concepts", concepts that are regulated by rules in the sense that instances of them can be brought about, have consequences and be terminated according to those rules:

> The term "institution of law", as I shall use it, is therefore to be understood as signifying those legal concepts which are regulated by sets of institutive, consequential and terminative rules, with the effect that instances of them are properly said to exist over a period of time, from the occurrence of an institutive act or event until the occurrence of a terminative act or event (MacCormick, 1974: 53).

According to MacCormick, *institutive rules* are those that "lay down that on the occurrence of a certain (perhaps complex) act or event a specific instance of the institution in question comes into existence" (*ibid.* 52); *consequential* are those rules that provide for the consequences the existence of an instance of a given institution has. The existence of one instance of the institution in question is part of the operative facts of these rules. Lastly, rules are *terminative* when they provide for the termination of the particular instance of the institution under consideration (*ibid.* at 53).

Contrary to what the title of his article could make us believe, MacCormick claims that from the fact that legal concepts (or at least some of them) are "institutions" (and hence that the existence, effects and termination of instances of them are determined according to rules) it does not follow that *the law* itself is an institution: "there is an almost overwhelming temptation . . . to treat the concept "law" like the concept "contract" as denoting an institution which is defined and regulated by the relevant set of institutive consequential and regulative rules" (MacCormick, 1974: 57). This temptation must be resisted, because some legal rules elude this characterisation in terms of constitutive, consequential and terminative rules. It would, therefore, be incorrect to assume that all legal norms are "like statutes in that they can be conceived as existing "validly" in virtue of clearly statable institutive rules":

> It is at least contestable whether there are clear criteria for the existence of rules of common law. Some have indeed contended that it is a fallacy of positivism to suppose that the common law can be represented as a system or rules (MacCormick, 1974: 57).

MacCormick believes that legal norms can exist that "cannot be understood as being established in virtue of necessary or sufficient criteria of validity". This constitutes an objection to the claim that the law is an institution, at least if we accept his definition of institutions as "concepts regulated by some set of institutive, consequential and terminative rules".

But the consequences of this admission might be more important than MacCormick thinks. For consider: if there are legal norms that can validly exist without having been produced according to some institutive rule, then the existence, consequences and termination of those rules of law is not controlled by institutive, consequential, and terminative rules *alone* (this is MacCormick's concession). Because (assuming that) the institutive, terminative and consequential rules relating to the common law as a source of law do not render sufficient and necessary criteria of validity, then the common law cannot be an institution. But if that is the case, then *all* legal concepts MacCormick *is* willing to call "institutions" and whose institutive, consequential and terminative rules are (at least partly) to be found in the rules of the common law, cannot be institutions because of the very same reason, i.e. because those rules would not completely regulate the (creation, consequences and termination of instances of the) concept. Sometimes legal norms validly exist, though no institutive rule has been followed to bring them into existence. But very much the same happens concerning not only legal norms, but also instances of what MacCormick does want to call "institutions of law": a contract, for example, can exist even if the institutive rules have not been followed (*see* the example discussed by MacCormick, 1974: 68), and it can fail to exist even if the institutive rules have been followed (for an example, *see* MacCormick and Weinberger, 1986: 12) Thus it seems that either the law is an institution along with the others, or none of them is.

Later in his lecture, MacCormick returns to this subject. He accepts (as did Hart, 1948) that institutive, consequential and terminative rules are defeasible, with the consequence that they cannot specify necessary and sufficient conditions for the existence of an instance of an institution of law. Legal principles justify an "open-ended" list of exceptions, and this is "fatal" for any attempt to represent the institutive rules as stating sufficient and necessary conditions for "for valid adjudication by tribunals or whatever". Even if we were to write down a list of all the exceptions imposed by court in cases of a certain kind, "we could not be confident that we had succeeded in listing the sufficient conditions for validity of a determination or an act of delegated legislation or whatever" (MacCormick, 1974: 70).

And what he says here about institutive rules can equally be said "in relation to the other types of rules which I have mentioned, and indeed of "rules of law" generally" (1974: 73). What rules of law lay down are only "presumptively sufficient" conditions: if the law imposes certain requirements for the validity of an act in law, then the fact that an act of that sort complies with those requirements implies that the act in question "ought to be presumed to be valid unless it is challenged" (1974: 72), but challenged it might be, with the consequence that what appeared to be a clear instance of a valid act performed in accordance with clear and valid institutive rules might turn out to be invalid (*ibid.* at 72).

So the fact that institutive, consequential and terminative rules can be defeated in concrete cases does not by itself imply that the concepts they regulate are not institutions, because we can take those rules as stating

"presumptively sufficient" conditions. But if this argument can do the trick for legal concepts, I cannot see why it could not do it for the law itself. In both cases we would have institutive, consequential and regulative rules that specify what is presumptively the case; and in both cases this would not prevent instances of the "institution" (i.e. a particular contract or a particular legal norm) from validly existing, even though no institutive rule has been followed to produce it.

MacCormick would not be so easily persuaded: "we neither have criteria of validity for legal principles, nor therefore a distinction between valid and invalid principles of law" (1974: 73). Though it is possible to give an account of what makes true the statement "the principle 'no one may profit from his or her own wrong' is a principle of English law", how those conditions actually work is something that cannot be understood without considering the values and purposes of the law. And to consider the values and purposes of the law is to consider the values and purposes the participants to a legal practice ascribe to them: "rules do not themselves have purposes, except in the sense that people may ascribe purposes to them" (MacCormick, 1974: 74).

The legal philosopher, according to MacCormick, has to recognise at this point that the explanation that is needed is not philosophical but sociological: "the philosopher may still pose questions, but he will have either to become a sociologist to answer some of them, or alternatively, have to wait for his sociological colleagues to give him the answers" (*ibid.*).

My objection to MacCormick's solution (treating legal concepts but not the law as an institution) is this: the lack of criteria for the validity of legal principles implies, up to the same extent, lack of criteria for the validity of instances of legal concepts like "contract" and the like. Because of that lack of criteria, we might be surprised to find that a given principle was part of the law though we did not know it. But (at least sometimes) the normative consequences of this "unexpected" (so to speak) principle will be to deny validity to some instance of the institution in question (e.g. a contract) that has been produced according to the relevant institutive rules (or, conversely, to lend validity to an instance that has not been so produced); hence insofar as we lack criteria for the validity of legal principles, we lack criteria for the validity of instances of "institutions of law"; insofar as the lack of those criteria is a reason for something *not to be* an institution, then neither the law nor contracts are institutions.

If, on the other hand, we follow MacCormick's advice and focus upon the fact that we *do* have presumptively sufficient conditions for the validity of instances of institutions of law, then could we not say that we also know what the presumptively sufficient conditions for the existence of legal principles are?

Notice again how all these complications would not in the least affect a theory of "football as institutional fact": the rule that specifies what a "goal" is does not specify "presumptively sufficient" conditions for something to be a goal, but necessary and sufficient conditions of anything to be one.

And here we can see how MacCormick's philosopher can go one step further: instead of taking it as a brute fact, she can try to explain what is it about the law that makes it so different from other normative systems in this regard. Such an explanation, we shall see, is partly empirical and partly conceptual. The argument in this chapter is (I hope) the beginning of it.[28]

TWO MODELS OF INSTITUTION

It is time to pull the threads of the argument together. To do this we can start with the distinction Searle failed to make between systems of rules (i.e. institutions) rather than rules. As was said before, this is a distinction between systems of rules (i.e. institutions) that constitute (i.e. create the possibility of institutional facts to be brought about) in order to produce some regulatory effect in the world (hence, as stipulated above, *regulatory* institutions) and systems of rules (i.e. institutions) that regulate some forms of behaviour in order to create the possibility of institutional facts to be brought about (hence *autonomous* institutions).

I think a distinction very much like the one I am trying to defend was in Wittgenstein's mind when he wrote

> Why don't I call cookery rules arbitrary, and why am I tempted to call the rules of grammar arbitrary? Because "cookery" is defined by its ends, whereas "speaking" is not. That is why the use of language is in a certain sense autonomous, as cooking and washing are not. You cook badly if you are guided in your cooking by rules other than right ones; but if you follow other rules than those of chess you are playing *another game*, and if you follow grammatical rules other than such-and-such ones, that does not mean you say something wrong, no, you are speaking of something else (1966: § 320).

[28] In his "The Epistemology of Judging" (1992), Thomas Morawetz criticises, in a way congenial to my own, the metaphor of games as "misleading" when applied to the law and other "deliberative practices" (of which he offers at 9 the following examples: "aesthetic debate, moral reasoning, historical discourse, and judicial decision-making"). But he does not offer an explanation for the fact that our deliberative practices are deliberative. The closest he gets to that is his remark that, in games, "the rules are fixed, and assumed to be known to all. But only the least important aspects of experience have this kind of simplicity. Only the least important aspects of life leave participants the option whether or not to play. In more immediate and important practices . . . we have a stake unavoidably and the shared rules-and-strategies are endlessly controversial" (Morawetz, 1992: 14–15).

This passage could be read as stating that non-deliberative practices are such either because they in some way "deal" with the "least important aspects of experience" or because they are "optional" in the sense that people can exercise an option not to play. Morawetz seems to believe that the latter is implied by the former, that is, that *because* non-deliberative practices deal with non-important aspects of experience, they allow for people to withdraw from them if they want. We have already seen that it is not "importance" that makes non-deliberative (in my terms, autonomous) practices non-deliberative (autonomous). As we shall see shortly below, Morawetz's second criterion (i.e. that those practices are in some way "optional") is, in my view, closer to the correct explanation.

The reason that, for participants, justifies the existence of an autonomous institution is the value they recognise in being able to engage in the particular kind of activity the institution sets up. Concerning institutions of this kind, it is pointless to look for an underlying activity the system is designated *to* regulate: it either does not exist, or, if it does, the point of the institution is not to regulate it, but to create a new, institutional thing *using it*. Of course, there are reasons why we want these new activities to exist but these are reasons for inventing the institution. It is not the case that we invent the institution because we want the underlying activities regulated in some particular ways: we do it because we want to be able to do something new, like playing football, or speaking a language, and so on (this is why for Wittgenstein these rules are, in a sense, "arbitrary" and speaking "autonomous").

Consider, for example the case of boxing. At first sight it might appear that rules of boxing *regulate* a fight in a way that is perfectly analogous to that in which the law regulates fights, but in the sense I have been using the expression, this is clearly inaccurate. The point of the institution of boxing is not to regulate fights (as, e.g., criminal law does), *but to create a new, institutional, form of fighting*. Of course, the creation of this institutional form of fighting called boxing is achieved (*inter alia*) by regulating the brute fact of a fight. But the *point* of (or the reason for) inventing the institution of boxing is not to regulate fights, but to *create the game*. Hence the rules are applicable only if you participate in the game, *because* you do so; if you are not boxing, then you are under no boxing obligation to apply the rules of boxing, even if you are a professional boxer (you might of course have some other reason for so doing: maybe you are better at fighting when you follow them, or you think that that is the only fair way of fighting, etc., but these are not counter-examples here).

Regulatory institutions are different: it is clearly wrong to say, regarding them, that we invent (say) the law because we want to create *ex-novo* new activities. Rather, we want to regulate in a certain way some pre-existing activities, actions, relationships, etc. (and in this sense the rules are, at least partially, "defined by their ends"). We want to be able not only to exchange goods, but also to have notions such as futurity and obligation linked to the exchange, because an exchange in these conditions (contract) seems to us more useful than a "brute" exchange (*see* Atiyah, 1982a: 1). Of course, to do this we have to invent institutional concepts like contract and the like, but the reason for so doing is our interest in the regulation of some forms of behaviour that exist outside the institution. Furthermore, it is not only not bizarre, but substantially accurate to say that because we want to regulate the killing of one human by another and economic transactions we have to invent the law.[29] Notice that if a given legal rule concerning an action φ (the celebration

[29] It goes without saying that the language I am using is in a sense particularly inaccurate: of course, "we" did not "invent" the concept of contract "because" we "wanted" such-and-such. The history of the emergence of legal institutions is more complex a subject. But I think that the argument stands any level of complexity in relation to that history, and so I am using this inaccurate language to facilitate the exposition.

of a contract, the transfer of property, etc.) exists, then you are under a legal obligation to apply that rule every time you do φ. But you do not have a (football-) obligation not to touch the ball with your hands if you are not playing football (*see* Rawls, 1955: 164).

As an illustration, recall Bankowski's explanation of the development of an institution (quoted above at 24): "the institution grows up on the back of the substantive reasons since *the reason that it is a promise can be seen as the universalisation of the substantive reason*" (1998, emphasis added).

In this model, institutions (like promising) are (and are *seen* as) universalisations of substantive reasons. The institution is not autonomous from the reasons for it. Notice, further, that what we say here of "institutions" could very well be applied to the rules of them: legal rules are seen as universalisation of substantive reasons, as "entrenched" generalisations (Schauer, 1991). This is the reason why, though the rule might (some would say, has to) have *some* insularity from those reasons, it cannot be completely cut off from them in the way the offside rule can. Therefore, if instead of trying to explain the emergence of an institution like promising we wanted to explain that of a game like football, or that of an institution of football like the penalty kick, we would find that Bankowski's interplay between the rule and the substantive reasons behind it is quite different. Granted, there is always a sense in which football grew on the back of substantive reasons, and to see this we can avail ourselves of the distinction between "to play" and "to play a game" (Opie and Opie, 1969: 2). Once upon a time, we can say, people did not play games, because no game had been invented. They only played. In some moment, one of the players told the others that it would be much more fun if they were to kick the ball through three posts instead of just among them. So they decided. Then other players noted that it would be even more fun if there were a limited pitch, and two teams with the same numbers of players, and so forth. Sooner or later they will start playing football, or some primitive form of it.

As we saw when discussing Searle's general theory, there is an asymmetry between regulatory and autonomous institutions here. When participants in a given social practice are evolving the institution of money, they need not be aware that they are imposing on whatever they are using as medium of exchange a meaning that is not exhausted by the physical properties of it. But nothing we could recognise as football can be played if we are not entitled to assume that the "players", in one form or another, are actually aware that those three posts at each end of the pitch have meaning in addition to their physical properties: on top of their being wooden posts, they are *goals*, and if the ball crosses them a point is scored. This cannot but be transparent to the players.

Because regarding regulatory practices participants need not be aware of the fact that they are evolving them, the interplay Bankowski sees between the rules and the reasons for them is quite different. This relationship in games is, I would argue, "one way only": because of the reasons discussed above (at 9f) FIFA decided to modify the offside rule. Once FIFA so decided, and only because of it, the new rule is a rule of football.[30] There is no going back to the reasons at the moment of applying the rule, as we have seen. Because of this, a "genetic" account of the emergence of football along Bankowskian lines might be of interest for the historian of football: it shows how football was brought about. But it would not help a referee who needs to apply the rules: compare the case of promising, in which such an account would indeed help someone who has to decide whether the fact that a friend is ill is relevant to her decision to keep a promise to be somewhere else at the moment her friend needs her company. In other words, the interesting thing about Bankowski's explanation of the emergence of moral or legal institutions is that it illuminates the interplay between the rules and the substantive reasons they are supposed to advance, interplay that is in turn explained because Bankowski shows the rules as "universalisation of substantive reasons". In the case of games there is no such interplay because the rules, though they might be universalisation of substantive reasons, are not to be *seen* as such by participants. They are seen as "simply what we do" when we play football.

This last point is important because it shows why I do not have to deny that there are substantive reasons for the rules of autonomous institutions (hence they need not be wholly "arbitrary"). Imagine that we are in a convention inventing a game. We can decide, for example, that we want a game of physical ability. That would rule out any game like chess or bridge. Furthermore, we can also decide that our game is to be one of team work, so tennis is excluded, and so forth; progressively, we write down the rules of football. We might decide that we want to allow *any* physical ability, including the ability to injure the adversary if this is useful. Or we can take a more sensible approach, and decide that we do not want to allow any move that can affect the physical integrity of any player. Once we have decided that, we need to introduce the pertinent rules: even in the first case, we will have to forbid the use of weapons (at least those which do not require the exercise of some physical ability). Furthermore, we could find that we want to make the game safer, and to punish any move that can be dangerous for a player. We shall find that there are two ways of achieving this aim (Hart, 1994: 125f): we can either elaborate a list of the moves we consider dangerous, or we can give the referee discretion to determine if a given move is dangerous (of course, we

[30] Throughout, I have been referring to FIFA as football's legislative body. This is, needless to say, for ease of exposition only. Of course, many people play football without even knowing in detail the rules approved by FIFA. This, if anything, makes my argument stronger: most of the time people need not agree in advance to the rules they are going to apply. They simply rely on their knowledge of the rules. And even in these circumstances it is most uncommon to find players disagreeing as to what the rules are or how they are to be applied.

can mix these two approaches up: this is what happens today in football). We will have to decide if we want more safety at the price of discretion, or if instead we want to deny umpires discretion at the price of some safety. My point is that, given that we are deciding how to build an autonomous institution, *everything is up for grabs*, though after each decision our space of *manoeuvre* will be smaller. At the very beginning, when we decided to invent a new game, every conceivable game was the possible outcome of our convention. After our first decision, as we saw, games like chess were ruled out; after our second, tennis and the like were. The point is that many of our decisions may be fully justified by substantive reasons: they need not be "arbitrary". What makes the rules we create "arbitrary" in Wittgenstein's sense is that those rules are to be seen by players as "simply what we do" and *not* as universalisations of substantive reasons (recall the case above at 10 of the referee applying the new offside rule.)

But a legal system is in this respect different from games, and to see this it could be useful to use an example here. Recall Fuller's case of the two men sleeping at the station. If we wanted to give hard cases in law the same explanation we gave to the hard case of Ronald Fuentes' handball (that is, an explanation based on the general defeasibility of concepts), we would have to say that, insofar as the first man was doing something that it would be non-controversial to classify as a "core" instance of the word "sleeping", he (and not the second, who wasn't sleeping) must be fined. But this solution would strike any sensible lawyer (and lay persons as well) as, at best, odd. If the first man is to be acquitted, however, this is not because we can say that he was not "really" sleeping, but because we think that the rule *should not be applied* to this case. The rule is to be seen as the universalisation of substantive reasons, and the point of Fuller's example is that any participant would immediately see that no substantive reason is served by fining the first man. Now imagine that the rule is not a legal one, but the rule of a peculiar game called "staying awake in railway stations". The game consists in avoidance of falling asleep in the station, and if you do you lose five points. Here the rule is not to be seen as such a universalisation, and as a consequence of that participants can agree that the first man must pay but the second should not, *if they are playing this peculiar game*.

In other words, insistence upon *indeterminacy of meaning* as the master explanation of legal disagreement is clearly insufficient. The problem is not that we are not sure about whether the first man was or was not sleeping in the station (because his was a "penumbral" instance of "sleeping"): we know he was (anyone who is not sure has to look up "sleeping" in the OED). The problem is, rather, that we are unsure that the rule should be applied to *this particular* case to the exclusion of all other considerations, though their explicit operative facts are indeed fulfilled.

And here we can go back to the point made before, about games being somewhat less serious or important than the law. As should be remembered,

this explanation was rejected because some games can be much more serious, for participants and spectators, than many legal disputes. So the point cannot be one about seriousness *simpliciter*. But if we read it in the light of the distinction between autonomous and regulatory institutions drawn above, we can reformulate it. We can say that the point of some institutions is to invent new activities while the point of others is to select, from a vast array of ways in which things can be done, those which are to be preferred. In the first case, then, the decision to participate in the activity *amounts to* a decision to abide by the rules, i.e. *not to question the application of the rules to particular cases*. If you do question that point, you fail to participate. A football player who thinks it is better to score goals with the hands will not be allowed to do so under present-day football regulations. Imagine him saying: "the point of football is to create a challenging game. If I am allowed to score with my hands, it will be more challenging than it actually is", and then going on to do it, Maradona-like. The relevant football rule should be applied, and the goal should be invalidated. His insistence on the validation of the goal for the reasons given will be taken as a signal that he did not really want to play football, but to invent another game (Rawls, 1955: 164). And if he is allowed to do so, nothing happens, except that the whole group starts playing a new game, certainly not football (some people like to say that this is how Rugby was invented). In this context, the most "serious" thing that can happen is that these people fail to play football. But there is nothing sacred (usually) about football, so they could perfectly well say, "yes, we are not playing football: we prefer to play this new game, rugby". It is in *this* sense that we can say that there is nothing *serious* about games: we can always decide to play another game.

It is for this reason that in autonomous institutions it sometimes appears that the "normative becomes, in a certain sense, descriptive" (Bankowski, 1996: 33). The rules are binding insofar as you want to participate in the activity. If you don't, the rules don't matter. Hence, the rules of an autonomous institution can be seen as descriptions of how you should behave if you want to play the game.

If the argument so far is correct, then all the considerations made about games can be applied to other institutions whose point is to invent a new activity. Consider, for example, the distinction between the rules of grammar and those of games, on the one hand, and those of style and of *fair play*, on the other (I am not implying that regulatory and autonomous institutions always come in pairs). To create the possibility of speaking English or of playing football, we need the rules of English grammar and those of football respectively. Before these rules are invented it is impossible to do one thing or the other. These rules do not exist in order to regulate the sounds or marks we produce, nor the activity of running around a ball (though indeed, in a sense, they do precisely that), but to create the very possibility of speaking (English) and playing (football). But *once* they have been created, *then* we can treat

these activities as pre-existing for another purpose, and so we can think that, *given that* we can speak English or play football, we want to do so in special ways: we want to speak beautifully, or to play in an elegant and sporting manner. This is the context for the emergence of a regulatory institution: now we need a set of rules to regulate the activities of speaking and playing (i.e. norms that single out *some* of many alternative ways of speaking and playing as preferable). In other words, when we are trying not to set up the activity, but to establish normative standards for the *better* way to do something we can do anyway (like speaking English or playing football), we leave the autonomous and enter into the regulatory model.[31] And, correspondingly, we lose the certainty the rules had in the former: now it is not beyond plausible contestation what the standards of style or *fair play* demand, since now rules are (and are seen as) universalisations of substantive reasons. It is important to notice here that the rules of style and those of *fair play* are clearly not rules of language/football: you don't have to master the rules of English style/*fair play* to be able to speak English/play football, though of course your speaking/playing will be better if you do (Marc Overmars' goal is a splendid proof of that: *see* above at 11, n.8) They exist precisely because it is possible to participate in the activity of speaking English or playing football in different ways, and their point is to signal some of these ways as better than others.

THE WEIGHTIER MATTERS OF THE LAW

The distinction I have drawn above is not an empirical one: it purports to be a conceptual one, between two different kinds of institution. But the fact that the difference is conceptual does not mean that the law is, as a matter of conceptual truth, necessarily "regulatory". The model a given institution belongs to is an empirical question (though not the distinction itself), one that is settled by the way the participants understand their institution.

Consider the following analogy: it is a matter of conceptual truth (i.e. something that is settled by the concept of "mode of production") that a mode of production includes humans, raw materials and means of production. Whether a particular mode of production is capitalist or feudal or something else is an empirical question, i.e. one that is settled by the kind of production relations that actually obtains in a particular society. But *given* that a mode of production is capitalist, it is a matter of conceptual truth that, *inter alia*, proletarians are formally free; similarly with the law. I would not object if someone were to claim that the law is what I call "regulatory" as a matter of conceptual truth. This would amount to a verbal stipulation concerning the meaning of

[31] I am aware that I am stretching the meaning of the word "institution" when I say that fair play and style are institutions. The emphasis here is to be placed on the "regulatory" bit. The word "institution" could be replaced, here and elsewhere, by the expression "normative system", "practice" or the like.

the word "law", and as such would be both unobjectionable and quite unhelpful. I would prefer to claim that this is an empirical question (indeed, we shall be looking here at concrete instances of legal practices I would like to call "autonomous"). But if the answer to this empirical question were to be that a given legal system is a regulatory institution, then some consequences would conceptually follow, consequences that explain why this distinction is important.

So let us consider what a legal system conceived of as an autonomous institution would be like. The point I want to make is nicely illustrated by the way in which *formalities* can be regarded in different legal cultures. Though any formality could be used here, I want to focus particularly on the formalities required for the validity of a contract.

It seems to us completely obvious that formalities are required for some *reason*, a reason that is related to the act to which the formality is attached (in other words we are used to seeing rules requiring formalities for the validity or enforceability of a contract as "the universalisation of some substantive reason": regulatory institutions). The contract of guarantee, for example is (was) considered particularly liable to be agreed between parties of unequal bargaining power, so if the contract rerquires to be in writing the weaker party will be in a better position to counter that inequality than if it is oral. So the (English) law requires the contract of guarantee to be in writing (this is Atiyah's explanation: *cf.* 1995: 164). The formality is required because some reason of substance suggests the convenience of its existence.

This way of looking at formalities is nowadays commonsensical: "insistence on form is widely thought by lawyers to be characteristic of primitive and less well-developed legal systems" (Atiyah, 1995: 163). But the question is, *why* are (so-called) primitive legal systems more rigorously formalistic? The thesis I want to entertain here is that law has not always been regarded as a (to put it in my words) regulatory institution. The insistence upon formalities, in a way that seems so bizarre to us, is one possible consequence of the law being understood as an autonomous institution.

Here we would have to imagine a society in which officials and subjects understand the law (and the world) in ways very different to our own. We would have to imagine a society in which the law is seen not as an *instrument* used to regulate social interaction, but as a technique that rests upon regularities that pertain to the very fabric of the world, very much like the way we understand the technique of bridge-building (or cookery). They would think of "obligation" as meaning literally a (quasi-) physical bond, a bond that can only be brought about following a predetermined procedure, in very much the same way in which we take a bridge (or a prawn cocktail) to be a physical thing that can be brought about following a predeterminate set of technical rules.

This is not a purely fantastic idea. Indeed, something like this is what the ancient Romans seem to have believed, as Reinhard Zimmermann has claimed

(1990: 1f, 82f). To be able to put someone under an obligation would then mean to be able to create such a bond. This bond is thought of neither in the way in which *we* think of an obligation, nor as a relation of pure power. Precisely through the use of such formalities "the creditor's real power over the body of the person who was liable came to be replaced by a magical power over him, and it was for this purpose that a formal ritual had to be performed" (Zimmermann, 1990: 83). The function of the formalities of the *stipulatio*,[32] in this context, is quite different to the functions we are used to thinking the formalities perform. We are used to seeing formalities as protecting or promoting some value, interest etc. (i.e. as the universalisation of some substantive reason): the interest of third parties which can be affected by the transaction, the interest of the weaker party before that of the stronger, the facilitation of proof, hence the possibility of having less and cheaper controversies, and soon. But for (ancient) Roman lawyers, according to Zimmermann, all of this was beside the point. The ritual was not required for policy-based considerations, but simply because that was the *only way* of getting things done:

> it was only by means of these rituals that legal transactions could be effected: compliance with the ritual formalities brought about a *real* (but invisible and in so far magical) *change* in the relationships between the parties concerned. The slightest mistake would wreck the whole transaction: every reader of fairy tales knows that magical effects can be engendered only by a most punctilious recital of a set formula . . .. The *actual reason* for the desired legal result was not the consent between the parties but the formal exchange of the words (Zimmermann, 1990: 83–4; emphasis added).[33]

It is not part of my argument that a formalistic understanding of formalities is only possible in autonomous institutions. Some legal formalities, in some legal systems, are nowadays thought of in a very formalistic way indeed. But the formality of these areas of modern law is based upon considerations of policy: they are (seen as) universalisations of substantive reasons. Hence it is always possible, at least in special cases, to go back to the raw "policy question"—and how "special" a case has to be is a *substantive* matter, i.e. something to be decided in the light of the policy-reasons underlying the formality (more on this later). In (ancient) Roman law, on the other hand, there was no "raw" moral (policy) question to go back to: the formalities were not required for substantive considerations, but because that was the only way in which a given effect could be brought about. This has as a consequence that the application of the rules becomes highly mechanical:

[32] The *stipulatio* was one of the most important contractual forms in Roman law. It was defined only by its form. Any obligation could be created using it (*see* Zimmermann, 1990: 68ff).

[33] The issue of the magical character of law in Rome is a controversial one: compare Hägerström (1953: 56ff), for whom the role of magic in Roman law was ubiquitous, with G. MacCormack's criticism of this thesis (MacCormack, 1969). The magic character of *ancient* Roman law is less controversial, though (*see* Kaser, 1967: 133).

> The most characteristic feature of archaic Roman jurisprudence is its tendency to endow every (sacral and) legal act with a definite form. Specific rituals had to be meticulously performed, precisely set forms of words to be uttered with great punctiliousness. The smallest mistake, a cough or a stutter, the use of a wrong term invalidated the whole act. This actional formalism corresponded to a similarly strict formalism in the interpretation of those ancient legal acts. No regard was had to the intention of the parties; what mattered were the *verba* used by them. The more rigid the interpretation, the more care was, in turn, bestowed on the formulation of the formulae. The drafters had try to eliminate every risk of ambiguity. This led to scrupulous attention to detail [and] to cumbrous enumerations . . .. Anyone who failed to employ such devices ran the risk of having to face unwelcome and unexpected consequences: as was experienced, for instance, by those who had taken the vow to sacrifice "*cuaequmque proximo vere nata essent apud se animalia*" ("whichever animated things were born in their house next spring"). Not only animals but their own children also were taken to be covered by these words (Zimmermann, 1990: 623).

To have an idea of what the law would be like in this context, we could well follow Zimmermann's advice and think of fairy tales: if you don't say the magic formula exactly as it should be said, you fail to produce the results you were looking for. Elmer's case would not have been a problem in this setting: it does not matter who (and for what reasons) gets the magic lamp, the genie will obey. In the terms of the argument presented here, there is no space for more or less reasonable interpretations of what the formalities are: interpretations are either correct or not (more strictly, one interpretation is correct and all the rest are not): *qui cadit a syllaba, cadit a causa.*

If we are to accept Zimmermann's claim about Roman law, my contention is that for ancient Roman lawyers the law was not regarded as anything like a social technique "to induce human beings—by means of the notion of this evil threatening them if they behave in a certain way, opposite to what is desired—to behave in the desired way" (Kelsen, 1934: 29) or the "enterprise of subjecting human conduct to the governance of rules" (Fuller, 1964: 106), but as a magical language that had to be mastered if some effects were to be produced, magical language that was created by the Gods and communicated to humans by priests—remember that in ancient Rome the law was administered by the Roman pontifices, who were state priests.[34] Note that there is no need for justification in this legal system: imagine one Roman farmer asking his lawyer: "why should I answer precisely '*spondeo*' to celebrate a *stipulatio*? Is it not enough to manifest my consent in any appropriate way?" The lawyer would say: "you simply cannot do otherwise if you want to celebrate such a contract". The situation is entirely similar to that of a child asking "why

[34] The main point holds even is this claim if historically false, i.e. even if the ancient Romans did not see their law as something given by the Gods: we do not think (not all of us, at least) that the laws of gravity were given by God, and that is not an obstacle to our conceiving of bridge-building as a technique that rests upon the basic structure of the world. In other words, how and why the participants came to view the world as they actually do is immaterial here.

cannot I move the king more than one space at a time?" or a naïve engineer asking "why should I build bridges in this particular way?" ("because if you don't you'll fail to play chess/build a bridge").

If I am right in this regard, then we should expect to find a different conception of legislation. Since the formalities for the validity of contracts were not universalisation of substantive reasons, policy considerations did not have any bearing on the selections of the specific forms required, nor upon the consequences of failing to follow them.

And we do find, for example, that though the Romans did have statutes forbidding the conclusion or the content of certain contracts, they used a system of statutory prohibitions that seems very. peculiar to the modern observer:

> Three different types of statutes were distinguished . . .: *leges imperfectae, leges minus quam perfectae* and *leges perfectae*. Only acts performed in violation of *leges perfectae* were void. *Leges minus quam perfectae* threatened the violator with a penalty, but did not invalidate the act itself. Infringement of a *lex imperfecta* led neither to a penalty nor to invalidity (Zimmermann, 1990: 697–8).

The question presents itself immediately: what was the point of *leges minus quam perfectae* and *imperfectae*? If the contract was to be forbidden, why not to use a *lex perfecta*? The answer becomes clear if we take into account that "in the early days of Roman law the validity of a transaction seems to have been judged only from the point of view of the required form". A contract was defined by its form, and if the forms had been fulfilled it was simply not possible to go back and invalidate it: "that statutory prohibitions could interfere with and indeed completely invalidate formal private acts was inconceivable to the lawyers and the law-makers of the earlier Republic" (Zimmermann, 1990: 698; *see*, for a different explanation Stein, 1966).

A similar point has been made by David Daube from a different perspective. Daube was intrigued by the peculiar verbal forms Romans used, and by how those forms changed during the centuries. Roman statutes usually contained the imperative form ("shall", "shall not"); in some of them, however, the imperative form is replaced by phrases like "it is needful", "it is proper" etc. Daube focuses upon the different meaning of verbal forms of the following kind: "'if anyone damages another's property, it will be needful for him to pay' and 'if anyone damages another's property, he shall be bound to pay'" (1956: 4). According to Daube, phrases of the former type

> express, not a direct command—"I order you to do this or that"—and not even a freely formed opinion—"In my view you should do this or that"—but a reference to some higher authority—"There are compelling reasons to do this or that" (*ibid.* 8).

So the reason why these verbal forms are so common in Roman law was, according to Daube, that the legislator did not see himself as *creating* the obligation to pay damages (to use the former example). It would be odd for us to say: "if anyone wants to build a bridge (or to prepare a prawn cocktail), he

or she is bound to . . ." rather than "if anyone wants to build a bridge (prepare a prawn cocktail), it will be needful for him or her to . . .". Hence, it should not be surprising by now to find out that these special verbal forms

> belong without exception to the sacred law. "If a man is killed by lightning . . ., it is not permissible to celebrate the funeral rites for him": evidently, this is not the decree of a free lawgiver, a lawgiver who might, if he liked, enjoin the opposite. It is, essentially, interpretation; It is the wise men's reading of the divine will. The priests . . . do not dictate to you. They inform you of the results of their studies of sacred things (Daube, 1956: 9).

The interesting point is that this indirect imperative form ("it is proper to . . .") is not used by republican and classical lawyers to speak of the requirements of the praetorian law (*ius honorarium*), but only to the old *ius civile*. Only the *ius civile* was understood in the magical sense with which I am now concerned: "a praetorian or aedilician obligation cannot be inferred from a search into the law or legally recognised transactions" (Daube, 1956: 15): it rests only upon the praetor's (aedile's) authority. So Daube's remarks lend support to Zimmermann's view: the law (i.e., the old *ius civile*) was not seen as a social institution created by humans to regulate their affairs, but as something that was part of the very structure of the world, something that could be mastered and put to use by humans if only they came to know it.

This is why "insistence on form is widely thought by lawyers to be characteristic of primitive and less well-developed legal systems" (Atiyah, 1995: 162): it reflects this kind of "magical" view of the law. Insistence on form, just for form's sake, is demonstrative of an understanding of law in which it is given, not made (just as the rules of chess are given to the players, not made by them). This attitude changes according to changes in the respective legal culture: "the attitude of a legal culture towards form reflects its self-image and maturity" (Zimmermann, 1990: 88). The important point is that when this attitude towards form *has* changed, controversies can arise. All formalities nowadays have been introduced to achieve some legislative purpose. That purpose, however, might be realised in different ways, even if formalities are not complied with. In these cases, "the sanction of invalidity therefore seems to overshoot the mark: it is not demanded by the policy underlying the rules requiring formality of the act . . . Equitable inroads have therefore from time to time been made in the domain of statutory forms (Zimmermann, *ibid.* 1990: 86–7).[35]

The transition from an autonomous to a regulatory conception of the law can also be seen in the Bible. Isaac's blessing of Jacob instead of Esau was valid, even though it was obtained with deceit (Gen. 27: 18–40): Jacob disguised himself as his brother and made his half-blind father believe he was

[35] *Cf.* Zimmermann (1990: 118–19), for the same point concerning the *sponsio* (suretyship): once ritual requirements have been relaxed (because there is no magic in them) "intricate problems of interpretation could arise" (concerning the *unitas actus*—the requirement of the *sponsio* to follow immediately the celebration of the respective *stipulatio*).

Esau. Because of his father's mistake, he received the blessing though he was not the first-born. The case is revealing because Esau "wept bitterly": "Bless me, too, my father", but there was no "unpicking" of the blessing, no "going back" to the raw substantive question, because there was no substantive question to go back to: "your brother came full of deceit and took your blessing" (Gen.: 35). The blessing was a (quasi-) physical thing, something Jacob had *taken* from Isaac though the former was not, under the law, entitled to it. The deception, having succeeded, could not affect the *fact* that Isaac did not have a blessing for Esau other than "by thy sword shalt thou live, and shalt serve thy brother" (Gen. 27: 40): if someone takes the prawn cocktail I am about to eat, I cannot eat it or give it to you, since I do not have it any more; Jacob took the blessing and Isaac could not go back and "invalidate" it, any more than I could "go back" and "invalidate" someone's eating my prawn cocktail and then eat it up myself.

David Daube has pointed out that "four at least of the tales of Jacob culminate in the appeal by the subtler disputant to those rigid, formalistic principles which can so often be found governing the legal or religious transactions of ancient peoples" (1944). However, he ends this piece with "a word of warning":

> no greater mistake can be made than to argue that, since the narrative here reviewed invariably leads up to the triumph of the party abusing certain formalistic principles of law, the characters described, and even the authors of the descriptions, must have been primitive men who did not see the flaws in their system. The exact opposite is true (Daube, 1944: 75).

According to Daube, "the proper question for us to pose is not, 'why did they not see that there might be an alternative to that strict, pedantic kind of law?' . . . but 'why did they apply, in some branches of the law at least, those strict, formalistic principles although they were fully aware of the possibility of unjust results?'" (*ibid.* at 75). In Chapter 6 we shall see that failure to answer the second of Daube's questions would prevent us from fully understanding Roman law. We shall also see that an answer to the first question will constitute a significant step towards answering the second. But Daube does not offer reasons for his reluctance to ask the first question. If we think of the villain's getting hold of the magic lamp in the story of Aladdin, we see that the law cannot be offered as an "alternative" to the rules governing the obedience of genies. Indeed, this might provide a clue as to the correct answer of the second question: maybe they did not see an alternative because that was simply the way the world was. "Unjust results" were simply a sad consequences of the way in which the world was ordered. What Daube considers to be obvious (that the problem of *verba* and *voluntas* as it appears in the law of contracts was "an alternative" to the rigid rules governing Isaac's blessing) was not necessarily obvious to the writer of Exodus.[36].

[36] Indeed, it is interesting to notice that Daube's arguments to show that the writer and contemporary readers of Jacob's story noticed the "flaws" in their system, that they were "wide

Be that as it may, notice the huge difference between this understanding of the law and Jesus's "new law". What was important was not the ritual fulfilment of the rules. The ruler of the Synagogue was indignant with Jesus for healing a possessed woman on the Sabbath: "there are six working days: come and be cured in one of them, and not on the Sabbath" (Luke 13: 14). This ritualistic way of understanding the law is scorned by Jesus: "here is this woman, a daughter of Abraham, who has been bound by Satan for eighteen long years: was it not right from her to be loosed from her bounds on the Sabbath?" (Luke 13: 16).

Jesus could be understood here as arguing that the substantive reason behind the law regarding the Sabbath was not affected by his healing of the woman. Since the law had to be seen as "universalisation of substantive reason", the law correctly understood was not an obstacle for his healing of the woman. But the ruler could have answered "if she has waited eighteen years, can't she wait one more day?: the law ought to be followed", implying that the law was not to be *seen* as universalisation of anything, but as "simply what we do". But he didn't: he was "covered with confusion while the mass of the people were delighted at all the wonderful things [Jesus] was doing" (Luke 13: 17).[37]

One could think from this that Jesus's law was not law at all, that his was a particularist ethics. But he clearly did not see his message in this way: "do not suppose that I have come to abolish the law and the prophets; I did not come to abolish, but to complete. Truly I tell you, so long as heaven and earth endure, not a letter, not a dot, will disappear from the law until all that must happen has happened" (Matt. 5: 17,18). Jesus's new law was regulatory law; an alternative translation of Matthew 5:18 in *The New English Bible* makes this point clearer: "Truly I tell you: so long as heaven and earth endure, not a letter, not a dot, will disappear from the law before all that it *stands for* is achieved" (emphasis added). So the message was precisely that the law was not a ritualistic–formalistic–magical set of rules that had to be fulfilled in detail (autonomous law), but something with a *point*, something that *stood for* something else (regulatory law). Later Jesus was to come back to this point:

> Alas for you, scribes and Pharisees, hypocrites! You pay tithes of mint and dill and cumin; but you have overlooked the weightier demands of the law—justice, mercy and good faith. It is these that you should have practised, without neglecting the others (Matt. 23: 23).

awake to the problem of *verba* and *voluntas* and similar difficulties" do not support his claim. The fact that these incidents "were taken note of and handed down" is, *pace* Daube, not "sufficient proof of this". The fact that we all know that Cinderella's spell was broken at midnight (though it would have been nice, wouldn't it, had it continued until the ball was over?) cannot be taken as proof that we all recognise the "flaws" in (that) system, for there is no system to begin with: Cinderella's world was one in which spells would last only till midnight, and that was it. Similarly, maybe the readers of Jacob's story saw in it simply a commentary on how useful it was to master the basic rules of the world.

[37] Jesus was constantly accused of not keeping the Sabbath: *cf.* John 7: 22–23; 9: 16, etc.

Commenting on this passage, Harold Berman has said "what the whole passage says is first, that the heart of the law is "justice and mercy and good faith", and second, that the lesser matters, the technicalities, the taxes, the "mint and anise and cumin" are also important, although they should be subordinated to the main purpose" (1993: 391).

In an autonomous institution the "mint and anise and cumin" is all that matters, insofar as you are participating in the activity the institution sets up (or, in autonomous institutions there is no distinction between the "mint and anise and cumin" and the "weightier matters"). But a regulatory institution is characterised by the fact that "justice and mercy and good faith" must be practised. This means that the "technicalities" *do* matter, but they are not (as in autonomous law) *all* that matter. Legal disagreement is explained by the fact that these two dimensions of regulatory law should be weighed up somehow: it is a disagreement about the correct way to balance them.

Jesus and the Pharisees would probably agree that they had to follow God's will. The difference was that the Pharisees believed that God's will *was* (as far as they could know it) the law. Hence, the law had to be followed blindly. To follow God's will was to follow God's law, because the law was the will: hence if you want to follow God's will, just follow the law; you need not ask what the law is really about, because it was given by God—he must know. But Jesus changed this: when the lawyer asked for a clear-cut definition (who is my neighbour?), he got a story and after that only the answer "go and do as he did" (Luke 10: 37).[38] Now to follow God's will (not only—or necessarily—the law in the formalistic–ritualistic view) was important: the (formalistically conceived) law was not enough. To the man who had followed the law since he was a boy, Jesus said: "one thing you lack; go, sell everything you have, and give to the poor, and you will have treasure in heaven. Then come and follow me" (Mark 10: 21).

Remember one of the characteristics of games, according to Huizinga:

> inside the play-ground an absolute and peculiar order reigns. Here we come across another, very positive feature of play: it creates order, is order. Into an imperfect world and into the confusion of life it brings a temporary, a limited perfection. Play demands order absolute and supreme. The least deviation from it "spoils the game", robs it from its character, and makes it worthless (1970: 29).

[38] Peter Winch has rightly emphasised that the parable of the Good Samaritan was offered as an explanation of what *the law* was. Jesus's first answer to the lawyer's question was : "what is written on the law?", and after the lawyer's answer, he said: "thou have answered right. This do, and thou shalt live". It was only when the lawyer, "willing to justify himself" asked Jesus about the *interpretation* of the law that "Jesus, answering, said, A certain man went down from Jerusalem to Jericho . . ." (Luke 10:26–30. *Cf.* Winch, 1987: 155f). Winch, furthermore, calls our attention to the fact that Jesus's answer to the lawyer's question was not linked to the latter's sharing a belief in God: "[the parable] did not appeal to the conception [of God as law-giver]: it challenges it. Or at least it commented on the conception in a way which presupposed that the moral modality to which the Samaritan responded would have a force for the parable's hearers *independently* of their commitment to any particular theological belief" (Winch, 1987: 160).

In this view, a religion (if understood as an autonomous institution) is not at all far from a game, particularly if we have in mind that in Huizinga's terms to play a game can be extremely "serious". Hence Huizinga's analysis of religion as a form of play. The same could be said of (ancient) Roman law: the least deviation from the wording of the *stipulatio* (for example, if the promissor answered the ritual question not with the word "spondeo", but with any other word, however similar or even identical in meaning) *made the whole thing worthless*. In classical Roman law (and even before) however, the raising of the *ius honorarium* and the *actiones bone fidei* changed this: it was no longer true that "the least deviation from it makes it worthless". Now some form of "justice, mercy and good faith" had to be done, without neglecting to pay "tithes of mint and dill and cumin": how these two things were served at the same time became debatable; hence the parties now had space for offering different views about what the law required, regardless of the words and the rituals used.

The distinction between two models of institution that has been put forward in this chapter is by no means new, though the labels I am using might be. Probably the clearest formulation of it, along with a realisation of its consequences for law and legal reasoning can be found in Max Weber's *Economy and Society*. For Weber, the formality of law meant that cases are decided on the basis of their "unambiguous general characteristics". Legal formalism, however, can be of different kinds, according to the nature of the general characteristics that are taken into account. On the one hand, they can be understood as "sense-data", like "the utterance of certain words, the execution of a signature, or the performance of a certain symbolic act with a fixed meaning". This is, for Weber, "the most rigorous type of legal formalism". It corresponds, if my argument is correct, to the formalism of ancient Roman law, and it is to be explained by the fact that for Romans the law was not seen as, in Pufendorf's words, "arising from imposition". The law, for example, makes the utterance of certain words (e.g. a matching word in the case of the Roman *stipulatio*: if the promissee asked, "do you *promise*?" the promissor had to answer "I *promise*") crucial to the celebration of a contract, so that if those words are not uttered there is no contract. This form of formalism, according to Weber, "exhausts itself in casuistry", because there is nothing beyond the actual fulfilment of the ritual requirement that is important for the existence of the contract. It exhausts itself in casuistry, I would argue, because no principle can control or defeat the application of the rule in the same way in which no principle of, say, "favour the most aggressive team" can control or defeat the application of the offside rule, regardless of the level of institutional support the rules of football could offer to a principle like the one mentioned.

Those "general characteristics", on the other hand, can be produced through "the logical analysis of meaning" so that "definitely fixed legal concepts in the form of highly abstract rules are formulated and applied".

This formalism views legal requirements (e.g. consent in the case of contracts) not simply taken as given, but as the product of logical rationality, because the law is indeed seen as "arising from imposition": there must be a *point* to the requirements of law, and that point is disclosed through "the logical analysis of meaning", leading to the formulation and application of "highly abstract rules". The reason why consent is important for a contract to be valid is no longer seen as inscrutable or irrelevant, hence if the promissor simply said "I do" rather than "I promise" it might well be acceptable. Thus, this second form of formalism "diminishes the significance of extrinsic elements and thus *softens the rigidity of concrete formalism*".

Interestingly enough, though Weber believed that this formalism of the "logical rationality" variety softens the rigidity of the first type of formalism, he also claimed that with it

> the contrast to "substantive rationality" is sharpened, because (substantive rationality) means that the decision of legal problems is influenced by norms different from those obtained through logical generalization of abstract interpretations of meaning.

In a system of substantive rationality, cases are solved according to, among others, "ethical imperatives, utilitarian and other expediential rules and political maxims". The formalism of the "logical rationality" variety is not substantive in this sense, because here logical rationality fulfils a "specifically systematic task", which is "the collection and rationalization by logical means of all the several rules recognized as legally valid into an internally consistent complex" (all quotations from Weber in this and the three previous paragraphs are from 1967: 61–2).[39]

Many of the themes to be developed in this book are present in Weber's remarks: law is *formal*, but this formality does not have to display the rigidity and formalism common in ancient legal systems. And yet the fact that law is not *that* formal does not imply that it dissolves into substantive reasoning. How the law can be formal-but-not-rigidly-formal and at the same time substantive-but-not-thoroughly-substantive is one of the major problems this book seeks to address.

For a completely different example, consider Hegel's criticism of mathematical knowledge (as discussed by Cohen, 1996). According to him, a mathematical explanation or proof is *external* to the subject, in so far as

> [t]he necessity does not arise from the nature of the theorem: it is imposed; and the injunction to draw just these lines, an infinite number of others being equally possible, is blindly acquiesced in, without our knowing anything further, except that, as we fondly believe, this will serve our purpose in producing the proof (Hegel, 1971: 102).

[39] Paul Amselek has also noticed that games are not analogous to law in important respects for reasons similar to those developed above. The difference, according to him, is that while in games "existence precedes essence" in law (and other institutions) essence precedes existence. *Cf.* (Amselek, 1988: 211). His remarks in this regard are, I believe, fully compatible with the argument offered in this chapter.

> *Qua* result the theorem is, no doubt, one that is *seen to be true*. But this eventuality has nothing to do with its content, but only with its relation to the knowing subject. The process of mathematical proof does not belong to the object; it is a function that takes place outside the matter at hand (*ibid.* 100–1, emphasis added).

Hegel's point, I believe, could be expressed as saying that the truth of a theorem is to be found in the definitions and rules of mathematics, not in "the object"—whatever that means. To understand a theorem is to be able to reproduce its demonstration: "if anyone came to know by measuring many right-angled triangles that their sides are related in the way everybody knows, we should regard knowledge so obtained as unsatisfactory" (Hegel, 1971: 100). But—though the result is, "no doubt" *seen to be true*—the knowing subject cannot see the *necessity* of the proof: "the proof takes a direction that begins *anywhere we like*, without our knowing as yet what relation this beginning has to the result to be brought out" (*ibid.* 102; emphasis added).

In other words, an institution like mathematics allows us to have absolute certainty with regard to mathematical knowledge, but this knowledge is in a way *defective*, because the process of proof is not internally related to the subject matter: we have to follow the process of proof in the hope that it will lead us to the demonstration we are seeking. The stages of that process are strictly determined by the (mathematical) rules, and the result is true in accordance to these rules.

The same happens, I would say, in any autonomous institution. Precisely because all that matters is the solution-according-to-the-institution, the process of finding it is *external* to the subject-matter. The justification of what (autonomous) law requires in these particular cases is *not* related in any way to the point at issue, but to the rules in question. As in mathematics according to Hegel, this understanding of law allows us to have absolute certainty about what it requires, but this absolute certainty has its price.

A SHORT PREVIEW

"What is law?" This is the central question of legal theory. "What is the law concerning this concrete case?" is the sort of question at which legal reasoning is addressed. After having shown that there is something about the nature of law as a social practice that makes the second question specially important, the chapters to come deal with the relation between these two questions. The thesis to be defended is that a theory of law implies a theory of legal reasoning; that is to say, that the second question is (at least partially) answered once the first question is answered. If this is correct, then we will be able to "read" a theory of legal reasoning into a theory of law and, reciprocally, to "read" a theory of law into a theoretical description of legal reasoning. The second central claim is that legal reasoning is formal (in the sense that it excludes substantive considerations) and yet substantive (in the sense that it cannot exclude all substantive considerations).

Chapters 2, 3 and 4 examine several traditional jurisprudential discussions in an attempt to demonstrate that a correct description of legal reasoning cannot show it to be completely formal (or exclusionary). This argument is used against the sources thesis (Chapter 2) to clear the way for an explanation of what makes legal reasoning theoretically interesting (Chapter 3). The argument presented in Chapter 3 is then used to rescue what I take to be the best reading of Lon L Fuller's original argument, both from Hart and contemporary Hartians and from Fuller himself (Chapter 4). Chapter 5 gives a fresh look to the issue of defeasibility in law and morality, trying to ascertain what are the "circumstances" of defeasibility, i.e. when and why does it make sense to say that a rule is defeasible.

Chapter 6 tries to offer an example of the argument so far by showing that we cannot really understand the law unless we understand legal reasoning. The example will be that of Roman law, and the claim will be that we cannot know what the correct solution was for a concrete case under Roman law just by learning about the Roman *rules*. Chapter 7 explores the way in which the argument concerning legal reasoning, as developed thus far, has consequences for a theory of law. Then four attempts to develop a theory of legal reasoning out of a theory of law that is, broadly speaking, positivistic, are shown to be defective. This is taken to be evidence of what is called a "tension" between legal reasoning and legal theory. Chapter 7 thus brings to completion the argument for the first thesis. Chapter 8 tries to solve the tension between a theory of law and legal reasoning diagnosed in Chapter 7 and completes the argument for the second thesis.