

# Rules, Reasons, and Acceptance

Adam Perry

Thesis for the degree of Doctor of Philosophy in Law

University of Oxford

Words: 75100

7 June 2012

# Abstract

---

In law as well as in ordinary life, it matters what rules societies have; but what does it mean for a society to have a rule? HLA Hart's famous answer is that for a society to have a rule is for there to be a certain social practice in that society, consisting of an external, behavioural aspect and an internal, attitudinal aspect. Hart's 'practice theory' dominates thinking in jurisprudence about social rules, but, I argue, there are serious problems with it. It would be better to adopt what I call the 'acceptance theory'.

In the early chapters of this thesis, I argue that the practice theory is both overinclusive and underinclusive. It is overinclusive because Hart's description of the 'internal aspect' is too general. It is underinclusive because the 'external aspect' is unnecessary. Once these criticisms are taken into account, what remains of the practice theory is the idea that a society has a rule because its members have a certain attitude. I spend much of this thesis determining the features of this attitude.

Ultimately, I focus on the attitude known as "acceptance" in the philosophy of action. Acceptance of a proposition simulates belief in that proposition, though it may be held independently of that belief. I argue that a person or society has a rule when that person, or that society's members, accept that some action ought to be performed, whatever their beliefs about the matter. This theory incorporates the plausible core of the practice theory, while avoiding its problems.

# Acknowledgements

---

I am very grateful to my supervisors for their support and guidance. At Oxford, I received supervision from John Broome, Leslie Green, and Grant Lamond. Scott Shapiro supervised me for a term while I was on exchange at Yale University. I also owe thanks to Timothy Endicott, Chris Essert, Matthew Kramer, Miriam Rodgers, Giovanni Sartor, Michael Sevel, and Jan van Zyl Smyt. I received financial support from the Commonwealth Scholarship Commission, the Government of British Columbia, the Law Foundation of British Columbia, and the Max Weber Programme at the European University Institute.

# Table of Contents

---

Table of Cases.....	6
Table of Statutes.....	7
Table of Figures.....	8
1 Introduction.....	9
1.1 A Persistent Question .....	9
1.2 Mandatory Rules .....	10
1.3 Social Rules .....	12
1.4 Rule-Possession.....	14
1.5 The Practice Theory .....	16
1.6 Two Approaches.....	18
1.7 The Acceptance Theory.....	19
1.8 Summary.....	20
2 Reasons.....	25
2.1 Introduction.....	25
2.2 Explanatory Reasons .....	26
2.3 Normative Reasons.....	27
2.4 Relations Between Explanatory and Normative Reasons .....	31
2.5. Conflicts Between Reasons .....	36
2.6 Conclusion .....	40
3 The External Aspect.....	42
3.1 The External Aspect: Introduction .....	42
3.2 Non-Compliance, Membership, and Evasion .....	45
3.3 Violations .....	52
3.4 Conclusion .....	55
4 The Internal Aspect.....	56
4.1 The Internal Aspect: Introduction .....	56
4.2 Raz's Summary .....	60
4.3 Warnock's Objection.....	63
4.4 MacCormick's Reply.....	66
4.5 Acceptance .....	68
4.6 The Internal Point of View .....	72
4.7 Conclusion .....	77

5 Strength and Strictness.....	78
5.1 Introduction.....	78
5.2 Strength .....	78
5.3 Strictness.....	88
5.4 Conclusion .....	94
6 Protected Reasons and Rule-Possession.....	95
6.1 Introduction.....	95
6.2 Protected Reasons and Rules .....	96
6.3 Rules and Partial Conflicts.....	103
6.4 Conclusion .....	116
7 Acceptance and Counterfactuals .....	118
7.1 Introduction.....	118
7.2 Counterfactuals and Reasons .....	119
7.3 Acceptance and Belief.....	132
7.4 Conclusion .....	146
8 Acceptance and Evidence Law.....	147
8.1 Introduction.....	147
8.2 Basics of Evidence Law .....	148
8.3 Exclusion.....	149
8.4 Presumptions and Fictions .....	152
8.5 Conclusion .....	162
9 Acceptance and Rule-Possession .....	164
9.1 Introduction.....	164
9.2 Content, Operation, Strength, Strictness .....	165
9.3 Reasons, Voluntariness, and Context-Dependence.....	179
9.4 The Acceptance Theory.....	189
9.5 Conclusion .....	194
10 Reasons, Possession, and Compliance .....	196
10.1 Introduction.....	196
10.2 Reasons to Have Rules.....	196
10.3 Reasons to Comply With Rules .....	204
10.4 Conclusion .....	207
Conclusion .....	208
Bibliography .....	211

# Table of Cases

---

<b>Case</b>	<b>Page</b>
<i>Campbell v Wallsend Slipway &amp; Engineering Co Ltd</i> [1978] ICR 1015 (DC)	157
<i>Chard v Chard</i> [1956] P 259 (DC)	156
<i>Goold v Evans &amp; Co</i> (1951) 2 TLR 1189 (CA)	148
<i>R v Reynolds</i> (1950) 1 KB 606 (CA)	149
<i>R v Lawson and others</i> (2007) 1 Cr App R 277 (CA)	150
<i>Woolmington v DPP</i> [1935] AC 462 (HL)	247
<i>Nicaragua v United States (The Merits)</i> [1986] ICJ 14	14
<i>North Sea Continental Shelf Case (Federal Republic of Germany v Denmark and the Netherlands)</i> [1969] ICJ 3	14
<i>Krulwitsch v United States</i> (1949) 336 US 440	151

# Table of Statutes

---

<b>Statute</b>	<b>Page</b>
Family Law Reform Act 1969	157
Law of Property Act 1925	152, 154
Homicide Act 1957	156
Police and Criminal Evidence Act 1984	149
Sexual Offences Act 2003	156
Town Police Clauses Act 1847	13

# Table of Figures

---

<b>Figure</b>	<b>Page</b>
1. Types of Counterfactual Reasons	123
2. Types of Acceptances	140
3. Typical Legal Presumptions	153
4. Types of Rules	168

# 1 Introduction

---

## 1.1 A PERSISTENT QUESTION

At the beginning of *The Concept of Law*, HLA Hart introduced three ‘persistent questions’ in jurisprudence. The first two are about obligation. Hart introduced the third this way:

At first sight it might seem that the statement that a legal system consists, in general at any rate, of *rules* could hardly be doubted or found difficult to understand. ... Yet dissatisfaction, confusion, and uncertainty concerning this seemingly unproblematic notion underlies much of the perplexity about the nature of law. What *are* rules? What does it mean to say that a rule *exists*? Do courts really apply rules or merely pretend to do so?<sup>1</sup>

After distinguishing mandatory and other types of rules, Hart returned to the existence of rules, reporting ‘the widest divergence of views as to the meaning of the assertion that a rule of this simple mandatory type exists. Some indeed find the notion utterly mysterious’.<sup>2</sup> Hart spent significant portions of *The Concept of Law* trying to dispel that mystery. His arguments were groundbreaking. Today they exert a huge influence on jurisprudence. However, there are problems with Hart’s account such that much of the mystery remains. In this thesis, I build on Hart’s work and try to add to our understanding of the existence of mandatory rules.

---

<sup>1</sup> HLA Hart, *The Concept Law* (2<sup>nd</sup> edn, OUP 1994) 8.

<sup>2</sup> Hart, *Concept of Law* (n 1) 93.

## 1.2 MANDATORY RULES

The word ‘rule’ has many senses. In one sense, a rule is a person’s or organization’s control or dominion over a people. In another sense, it is a description of the normal or customary state of things. In the sense relevant to this thesis, a rule is something that governs conduct. With respect to some conduct-governing rules, it makes sense to say the rule permits or allows conduct and, with respect to others, to say the rule mandates, requires, or prohibits conduct. These conduct-governing rules are respectively called ‘permissive’ and ‘mandatory’ rules.

Hart focused on the existence of mandatory rules, and I will do the same. My main reason for doing so is that the paradigmatic legal rule is a mandatory rule. Much of the criminal law, for instance, is made up of rules that prohibit conduct, including rules against assault and theft. Moreover, we can only fully understand other legal rules by explaining their relationship to mandatory rules. Mandatory rules are not the special province of law, however. Such rules exist in countless areas of life. There are mandatory rules of grammar, etiquette, games, religion, morality, and so on.

Mandatory rules can be expressed in words or not, but even when they are expressed in words, there is no special form for stating them. Max Black describes the variety of rule-formulations:

A request for an example of rule can be satisfied by producing a full sentence in the indicative (‘The dealer at bridge always bids first’), a ‘that’-clause (‘[The rule] that students enrolled in American colleges must have a high school diploma’), a verbal clause in the infinitive (‘To show small capitals in proof, underline twice’), a full sentence in the imperative (‘Do unto others as you would be done by’), an imperative phrase (‘No smoking in classrooms’) – and no doubt in many other ways.<sup>3</sup>

---

<sup>3</sup> Max Black, *Models and Metaphors* (Cornell University Press 1962) 106. See also, Georg Henrik von Wright, *Norm and Action* (Routledge & Kegan Paul 1963) 102: ‘[N]orm formulations, linguistically, are a very varied bunch. They cut across several grammatical types of sentences without including or being included in any one type. One must therefore warn against the idea of basing the conceptual study of norms on a logical study of certain linguistic forms of discourse’.

That there is no special formula for expressing a mandatory rule means that it is especially important not to equate the formulation of a rule with the rule itself. For example, a rule that can be formulated as ‘The dealer at bridge always bids first’ can be reformulated as ‘The dealer at bridge must bid first’. There is one rule, which can be formulated in two ways.

Although formulations of mandatory rules take different forms, it is always possible to state a mandatory rule in the form of an elementary ‘ought’ statement.<sup>4</sup> Schematically, any mandatory rule can be stated in the form ‘ $\mathcal{A}$ s ought to  $\phi$ ’, where  $\mathcal{A}$ s are members of a class of person  $\mathcal{A}$  and ‘ $\phi$ ’ is an act-type<sup>5</sup>. For example, ‘The dealer at bridge always bids first’ can be restated as ‘The dealer at bridge ought to bid first’. As Gideon Gottlieb writes, ‘any utterance which is designed to function as a rule *must* have the potential of being reduced, expanded, analyzed, or translated into [this] standard form’.<sup>6</sup> Statements of rules may not actually have this form. ‘The crucial question ... is whether it lends itself to a *restatement* in normative form’.<sup>7</sup>

When I use the term ‘rule’ in this thesis, I mean ‘mandatory rule’, unless I state otherwise.

---

<sup>4</sup> Joseph Raz, *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990) 50: ‘The content of a rule can be stated by the use of an elementary “ought” sentence to the effect that some person ought to perform a certain action’. See also: Black (n 3) 106-108; von Wright (n 3) 100-102; Gideon Gottlieb, *The Logic of Choice* (George Allen & Unwin 1968) 40; Neil MacCormick, *Legal Reasoning and Legal Theory* (OUP 1978) 43, 45; Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (OUP 1991) 23; Robert Alexy, *A Theory of Constitutional Rights* (Julian Rivers tr, OUP 2002) 22-25; William Twining and David Miers, *How to do Things with Rules* (5<sup>th</sup> edn, Cambridge University Press 2010) 141-142.

<sup>5</sup> It is sometimes claimed that rules only require generic actions. von Wright, for example, claims that ‘generality with respect to occasion’ is a necessary feature of rules, meaning that a rule-act must be capable of being performed on an unlimited number of occasions, over an unlimited time-span. See von Wright (n 3) 79-81, 83. von Wright does not argue for his view, and I do not see the motivation for it. It seems clear that some rules apply for a limited time-span, such as the *ad hoc* rules card players adopt for the duration of one hand (e.g., ‘Jacks are wild’). So while my argument is compatible with von Wright’s position, I do not assume it is correct. In this respect my approach is the same as Raz’s: see Raz, *Practical Reason and Norms* (n 4) 50. In any event, the issue does not seem to have any deep theoretical significance.

<sup>6</sup> Gottlieb, *Logic of Choice* (n 4) 40. In my schema I have included the circumstances in which the act is performed as part of the act-description. Gottlieb, like most scholars, prefers to state the two separately. Scholars who take this approach term the circumstances in which the act is required the rule’s *protasis*, *operative facts*, *factual predicate*, or *hypothesis*. The specification of the act that is required or permitted and by whom is the rule’s *apodisis* or *consequent*. See, for example: Schauer, *Playing by the Rules* (n 4) 23-24; MacCormick, *Legal Reasoning* (n 4) 43, 45; Twining and Miers, *How to do Things with Rules* (n 4) 140-144.

<sup>7</sup> Gottlieb, *Logic of Choice* (n 4) 40. Emphasis in original.

### 1.3 SOCIAL RULES

There are many questions we can ask about rules. We can ask: why have rules? And, why comply<sup>8</sup> with them? But first we need to ask, as Hart did, about the existence of rules. We need some idea of the conditions under which rules exist to know whether to have one or whether to comply with one. In this thesis, I start with the existence of rules, then broaden my discussion to include the justifications for having and complying with rules.<sup>9</sup>

Some facts are ‘social facts’, meaning facts about people’s actions or attitudes, whereas other facts are not. Some rules exist in virtue of social facts – Stephen Perry calls them ‘social artifacts’<sup>10</sup> – whereas other rules do not. Moral rules, for example, do not exist in virtue of social facts. There could be a moral rule against, say, lying whatever anyone does and whatever their attitudes are towards lying. On the other end of the spectrum, one of the tenets of legal positivism is that the existence of a law – and hence a legal rule – is exclusively a matter of social facts.<sup>11</sup> Joseph Raz writes: ‘In the most general terms the positivist social

---

<sup>8</sup> As I use the term, to “comply” with a mandatory rule is to do what it requires. I use “conform” in the same way. I provide a somewhat more precise definition in Section 3.1.

<sup>9</sup> There are other important questions, too. They include questions about how to interpret mandatory rules and how mandatory rules are related to other types of rules. Such questions are beyond the scope of this thesis.

<sup>10</sup> Stephen Perry, ‘Hart on Social Rules and the Foundations of Law: Liberating the Internal Point of View’ (2006) 75 *Fordham L Rev* 1171, 1147. Also 1175: ‘[A] norm is simply a certain kind of artifact, the existence of which depends on certain facts about human behaviour, attitudes, or beliefs’.

<sup>11</sup> For leading statements of the positivist claim, see: Gerald Postema, ‘Coordination at the Foundations of Law’ (1982) 11 *The Journal of Legal Studies* 165, 165: ‘Law is a social fact; what is and what is not to count as law is a matter of fact about human social behaviour and institutions ...’; David Lyons, *Ethics and the Rule of Law* (Cambridge University Press 1984) 37: ‘Laws are seen (by legal positivists) as brute social facts’; Ronald Dworkin, *A Matter of Principle* (Harvard University Press 1985) 131: ‘Legal positivism has many different forms, but they all have in common the idea that law exists only in virtue of some human act or decision’; Jules Coleman, *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory* (OUP 2003) 75: ‘Positivism claims that the possibility of legal authority is to be explained in terms of certain *social facts*. Call this the “social fact thesis”; no claim is more central to legal positivism’; Leslie Green, ‘Legal Positivism’, *The Stanford Encyclopedia of Philosophy* (2003) <<http://plato.stanford.edu/entries/legal-positivism/>> accessed 6 July 2011: ‘Legal positivism is the thesis that the existence and content of law depends on social facts and not on its merits’; Andrei Marmor, *Philosophy of Law* (Princeton University Press 2011) 4: ‘Legality is constituted by a complex set of facts relating to people’s actions, beliefs, and attitudes, and those social facts basically exhaust the conditions of legal validity’.

thesis is that what is law and what is not is a matter of social fact ...'.<sup>12</sup> Because legal positivism is the dominant position in legal philosophy, I will focus on rules whose existence is solely a matter of social facts.

It is sometimes clear what social facts give rise to a rule. As Geoffrey Warnock writes, '[s]ome persons and institutions have a more or less clear and uncontested authority to *make* rules; and in such cases there exist, there are, those rules that they have made, provided of course that they have not been subsequently abrogated ...'.<sup>13</sup> For a rule to exist, it is enough that it was properly made (laid down, enacted, etc.) and has not been abrogated. Nothing more is needed. For example, in the nineteenth century a law was enacted in Britain that prohibited anyone from flying a kite. Nowadays people in Britain regularly fly kites. That law is ignored and indeed most people do not even know it exists. Yet it does exist and it will continue to exist unless and until it is repealed.<sup>14</sup>

Other rules present greater difficulties. 'We must admit', Warnock writes, 'that there are rules which no rule-maker has ever made; and it is a separate question what the "existence" of these consists in'.<sup>15</sup> The kind of rule Warnock has in mind includes the rules of etiquette, grammar, manners, and 'social mores' generally. As Kurt Baier explains, '[r]ules of this sort have not been laid down by any one in particular nor are they modified or abrogated by specially authorized persons. Nevertheless they "hold" in certain groups and usually vary from one group to another'.<sup>16</sup> These rules exist in virtue of social facts (unlike the moral rule against lying) yet are unmade (unlike the legal rule against flying kites). They are called 'social rules'.

---

<sup>12</sup> Joseph Raz, *The Authority of Law* (OUP 1979) 37.

<sup>13</sup> Geoffrey Warnock, *The Object of Morality* (Methuen & Co Ltd 1971) 43.

<sup>14</sup> S 28, Town Police Clauses Act 1847. Sometimes a prescribed rule can cease to exist by desuetude. Whether this is true of this rule is not important for my discussion.

<sup>15</sup> Warnock, *Object of Morality* (n 13) 44.

<sup>16</sup> Kurt Baier, *The Moral Point of View: A Rational Basis for Ethics* (Cornell University Press 1958) 124.

Social rules of legal relevance include the rules of customary international law,<sup>17</sup> constitutional conventions<sup>18</sup>, and many others<sup>19</sup>. The most important social rule, for legal scholars, is a rule of recognition. Hart famously claimed that in every community that has a legal system there is a special type of rule – an ultimate rule of recognition – that determines the kinds of actions or events that create, modify, and abrogate law in that community.<sup>20</sup> In other words, the ultimate rule of recognition of a legal system determines what is a source of law in a community. Hart described it as a type of ‘judicial customary rule’<sup>21</sup>, clearly distinguishing it from deliberately created rules.

Social rules are both more fundamental and more puzzling than deliberately created rules. The interesting questions about deliberately created rules are about what gives someone the authority to create a rule. In the legal context, that shifts the focus to the relevant rules of recognition, which are not themselves deliberately created rules. Social rules are more puzzling because there is no obvious explanation for what makes it the case that there is such a rule. It is a matter of actions or attitudes, but it is unclear what actions, and what attitudes.

#### 1.4 RULE-POSSESSION

A social rule exists in virtue of facts about some society, and we indicate which society that is by saying that the rule exists ‘in’ that society or ‘among’ its members. The language of

---

<sup>17</sup> *North Sea Continental Shelf Case (Federal Republic of Germany v Denmark and the Netherlands)* (1969) ICJ 3; *Nicaragua v United States (The Merits)* (1986) ICJ 14.

<sup>18</sup> Joseph Jaconelli, ‘The Nature of Constitutional Conventions’ (1999) 19 *Legal Studies* 24. See also: Dennis Lloyd, *The Idea of Law* (Penguin Books 1964) 247-248; AD Woolley, ‘The Existence of Rules’ (1967) 1 *Nous* 63, 74-75.

<sup>19</sup> For many other examples of customary laws, see, David Bederman, *Custom as a Source of Law* (Cambridge University Press 2010) chs 5-9; Amanda Perreau-Saussine and James Bernard Murphy (eds), *The Nature of Customary Law* (Cambridge University Press 2007) chs 7,8, 12; Alison Dundes Renteln and Alan Dundes, *Folk Law: Essays in the Theory and Practice of Lex Non Scripta* (University of Wisconsin Press 1995) vol II.

<sup>20</sup> Hart’s account of rules of recognition is scattered throughout *The Concept of Law* (n 1). The most important passages for my purposes are 94-95 and 100-110.

<sup>21</sup> Hart, *The Concept of Law* (n 1) 256.

possession serves the same purpose: the rules that a society generates are ‘its’ rules; they are the rules that the society ‘has’. We say, for example, that there are rules regarding queuing in British society, or that British society has such rules, by which we mean that there are social rules regarding queuing because of how things are in British society.<sup>22</sup> (Although I mean to reflect ordinary usage, if you disagree, you should take my usage as stipulative. Nothing crucial turns on the terminology.<sup>23</sup>)

We usually care more about whether a rule exists in a certain society than about whether it exists *simpliciter*. People want to know what rules of etiquette and manners exist in their societies, or the societies of those with whom they interact, partly because it helps them to avoid the social opprobrium incurred by violating such rules. International lawyers and judges are not interested in whether a certain customary rule simply exists. They are interested in whether it exists among certain states, because that tells them about the duties and rights of those states under international law. Likewise, constitutional actors and scholars want to know what rules exist among constitutional actors, because that tells them about what actions are and are not constitutional.

When it comes time to address his ‘persistent question’ about the existence of rules, Hart focuses on the existence of a rule in a society.<sup>24</sup> I will do the same, the only difference being that I also consider what it is for an individual to have a rule as a way of shedding light on the social issue. Here, then, is the question I try to answer in this thesis: *what is it for a society*

---

<sup>22</sup> You might think that by saying that British society has rules regarding queuing we mean, in part, that there are rules regarding queuing that *apply* to members of British society. I give some counterexamples in Section 3.2.

<sup>23</sup> Theodore Benditt says that ‘[s]ocial rules are not the only kinds of rules that can in some sense exist in a society. For it is evident that at least some of the rules that are part of a legal system ... are not social rules’. *Law as Rule and Principle* (Stanford University Press 1978) 52. Benditt seems to find it natural to say of an enacted rule that it exists in some society. That may be true in special cases, but normally we say that such laws exist in *states* or other social institutions. It sounds more natural to say, for example, that there exists an income tax law in Canada, or, better still, that Canada *has* an income tax law, than to say that an income tax law exists in Canadian society.

<sup>24</sup> Hart, *Concept of Law* (n 1) 9, 55, 56, 255.

*or a person to have a rule?* (Sometimes, to save words, I will simply ask what it is to have a rule.)

The answer will take the form of a theory of ‘rule-possession’.

## 1.5 THE PRACTICE THEORY

According to Hart, for a society to have a rule, its members must generally do as that (putative<sup>25</sup>) rule requires. There must be a pattern or convergence of behaviour in that society. But the society’s members must do more. Rules are not generated by mere ‘habits’ or ‘customs’. We know that, says Hart, because ‘[m]ere convergence in behaviour between members of a social group may exist (all may regularly drink tea at breakfast or go weekly to the cinema) and yet there may be no rule *requiring* it’<sup>26</sup>. The second element is a certain ‘critical reflective attitude’ held by the members of the society in question. For those who possess this attitude, ‘deviations [from the rule] are generally regarded as lapses or faults open to criticism, and threatened deviations meet with pressure for conformity’<sup>27</sup>. Moreover, ‘[c]riticism for deviation is regarded as legitimate or justified ... as are demands for compliance with the standard when deviation is threatened’<sup>28</sup>. These demands, criticisms, and acknowledgements of legitimacy are expressed using normative vocabulary, including ‘ought’, ‘must’, ‘should’, ‘right, and ‘wrong’.<sup>29</sup>

Hart calls a pattern of compliance with a social rule the rule’s ‘external aspect’.<sup>30</sup> Possession of the right critical reflective attitude by members of the relevant society is the

---

<sup>25</sup> When I use phrases like this in the future, I leave it implicit that the rule is only putative.

<sup>26</sup> Hart, *Concept of Law* (n 1) 9.

<sup>27</sup> Hart, *Concept of Law* (n 1) 55.

<sup>28</sup> Hart, *Concept of Law* (n 1) 55-56.

<sup>29</sup> Hart, *Concept of Law* (n 1) 57. See also: HLA Hart, ‘Scandinavian Realism’ (1959) 17 Cambridge Law Journal 233, 238.

<sup>30</sup> Hart, *Concept of Law* (n 1) 56.

rule's 'internal aspect'.<sup>31</sup> Each aspect is necessary for a society to have a rule and together they are sufficient. (Although Hart frames the theory in terms of social rules, it can be adapted for personal rules.) Hart's theory has come to be known as the 'practice theory' because it treats the existence of a rule as a matter of a social practice consisting of the internal and external aspect.<sup>32</sup> Although Hart formulated the practice theory more than fifty years ago, in jurisprudence it remains the dominant theory of the existence of social rules. Nonetheless, there are serious problems with the practice theory.

One is that the theory is underinclusive. As we shall see below, contrary to what Hart claims, sometimes a society has a mandatory rule with which its members do not generally comply (e.g., rules that only apply to guests or visitors). Without its external aspect, the internal aspect of a rule takes on added importance. This leads to the second problem. When Hart described the critical reflective attitude possession of which is a social rule's internal aspect, he focused on how that attitude leads people to act. But, as Warnock pointed out, how it leads them to act is just how a belief that an action ought to be performed leads someone to act (e.g., to make criticisms and demands). The practice theory is therefore unable to distinguish cases in which a society has a rule from cases in which its members merely generally do something which they believe ought to be done.<sup>33</sup> The practice theory is not only underinclusive – it is overinclusive, too. These problems with the practice theory means that there is room – indeed, need – for an alternative theory of what it is for a society to have a rule.

---

<sup>31</sup> Hart, *Concept of Law* (n 1) 56.

<sup>32</sup> Hart, *Concept of Law* (n 1) 255.

<sup>33</sup> Warnock (n 13) 45-46.

## 1.6 TWO APPROACHES

Although we need to find an alternative to the practice theory, we do not need to start from scratch. Warnock's objection to the practice theory tells us that, when a society has a rule, its members act as if they believe that rule ought to be complied with. This is interesting because of what else it suggests. The natural explanation for why people act as if they believe something is that they actually believe it. It seems unobjectionable, then, to say:

- (1) If a society has a rule, its members believe it ought to be complied with.

There is something else we know about having a rule, which I shall demonstrate through examples. It is that, sometimes, a society has a rule but its members do not believe the 'balance of reasons' favours complying with it. (Indeed, sometimes, they believe the balance of reasons does *not* favour complying with it.) Now, the conventional view among moral philosophers is that what ought to be done is just what the balance of reasons favours. So, provided this view is correct, and that people who have a rule are not often and seriously conceptually confused, we can say:

- (2) Sometimes, a society has a rule, but its members do not believe it ought to be complied with.

Claims (1) and (2) are inconsistent, so we need to avoid committing ourselves to both of them.

Broadly speaking, there are two approaches we can take. The first approach is to accept (1) and reject (2). In order to reject (2), we need to show that the conventional view is incorrect and the balance of reasons does not always determine 'ought'. At this point in the discussion, Joseph Raz's famous account of protected reasons becomes highly relevant. According to Raz, if there is a protected reason for an action, it may be that this action ought

to be performed, though the balance of reasons does not favour it. Were we to accept this claim, we could develop a theory according to which possession of a rule is essentially a matter of the belief in a protected reason to comply with that rule. We lack a convincing argument for why protected reasons would have this special status, however, and without it the protected reasons-based theory of rule-possession is implausible. (At no point do I attribute this or any other theory of rule-possession to Raz himself; it is a theory that someone might try to develop using resources found in Raz's work on practical reason.)

The second approach is to accept (2) and reject (1). In order to reject (1), we need to identify an attitude that leads people to act as if they believe that something ought to be done, but that does not 'entail' that belief. Scholars who work on the philosophy of action and, to a lesser degree, the law of evidence make use of an attitude called "acceptance" that has these features. Acceptance of a proposition leads you to act in the same ways as a belief in that proposition normally would, but it does not entail that belief. Acceptance that a rule ought to be complied with thus leads people to act as if they believe it ought to be complied with, even though they may not actually believe that. The attitude has the right motivational force, and the right relationship to the relevant beliefs, making it a compelling basis for a theory of rule-possession.

## 1.7 THE ACCEPTANCE THEORY

I ultimately defend an acceptance-based theory of rule-possession, which I call the 'acceptance theory'. The acceptance theory says, roughly, that for a person or society to have a rule is for that person, or the members of that society, to be in a state of acceptance towards the proposition that this rule ought to be complied with. The main argument for this theory is that it can explain why people who have rules act as they do, without attributing to them

beliefs that they may lack. Another argument for this theory is that the features of acceptance other than its motivational force (for example, its responsiveness to practical reasons) are all features of the attitude on which rule-possession depends. (At least some of these features Hart described as features of his critical reflective attitude.) In the course of making these arguments, I also show that, by varying the content or scope of operation of an acceptance, we can explain the varying strengths and strictnesses of the rules people possess.

The acceptance theory is meant to be an answer to the most puzzling part of Hart's 'persistent question'. As such, it is meant to be an alternative to Hart's own answer, namely, the practice theory. But the acceptance theory also builds on the practice theory, and the similarities are as important as the differences. Given the importance of Hart's theory in jurisprudence, and the insight it offers into rule-possession, I regard these similarities as welcome. Indeed, one way to think of the acceptance theory is as an attempt to contribute to an answer to Hart's question along lines Hart himself may have been thinking.

## 1.8 SUMMARY

In this thesis I start with some basic topics in practical reason (chapter 2) and then proceed to a discussion of rule-possession. I begin with the practice theory (3 and 4) and an account of the varying strengths and strictnesses of mandatory rules (5). After considering a protected reasons-based theory of rule-possession (6), I move to acceptance. I introduce acceptance and reasons to act 'as if' (7), and demonstrate their relevance to the law of evidence (8). In the culmination of this thesis, I argue for the acceptance theory of rule-possession (9). I conclude by discussing some of the implications of the acceptance theory for the justifications for having and complying with rules (10).

## *Chapter 2 – Reasons*

To understand what it is to have a rule that something ought to be done, we need a rough account of reasons and, relatedly, what it means that something ought to be done. I borrow much of this account from Raz. I adopt, with slight modifications, Raz's accounts of explanatory and normative reasons (2.2. and 2.3). I describe some of the relations between explanatory and normative reasons, and set out Raz's account of 'exclusionary' and 'protected reasons' (2.4). I describe the varieties of conflicts between reasons and the relation between reasons and 'ought' (2.5).

## *Chapter 3 – The External Aspect*

With my discussion of reasons and 'ought' completed, I turn to rule-possession. I start my discussion as Hart did: with the idea that for a society to have a rule is simply for there to be a pattern of compliance with that rule in that society. Hart showed that a pattern of compliance was insufficient for a society to have a rule, but he thought social rules must still have an 'external aspect' (3.2). I go further and reject the idea in its entirety. I reject it on the basis of two types of examples. The first type of example is meant to show that a society may have a rule with which none of its members ever comply (3.3). The second is meant to show that a society may have a rule that its members generally violate (3.4).

## *Chapter 4 – The Internal Aspect*

Having discussed the 'external aspect' of social rules, I turn to their 'internal aspect'. I set out Hart's account of the internal aspect (4.1 and 4.2). I describe Warnock's objection that Hart's

account is overinclusive (4.3). I consider one response to this objection, suggested by Neil MacCormick, but reject it because it builds on the problematic external aspect (4.4). I describe some of Hart's remarks on 'acceptance', which are suggestive of a different type of response to Warnock's objection (4.5). I discuss Kevin Toh's well-known claim that Hart was what we would now call a 'norm-expressivist', and I explain why my project and Toh's are largely independent of each other (4.6).

### *Chapter 5 – Strength and Strictness*

By this point in the thesis I have claimed that for a society to have a rule is for its members to possess a certain attitude, one which leads them to act as if they believe that rule ought to be complied with. I explain that the 'ought' may be all-things-considered or relative (5.2). I show, too, that the attitude in question does not entail the belief that the balance of reasons favours complying with that rule. However, its motivational force may be absent when its possessor is sufficiently confident that, in a particular case, the balance of reasons does not favour compliance (5.3). Based on these points, I identify dimensions of strength and strictness along which rules may range, and give examples of four types of rule.

### *Chapter 6 – Protected Reasons and Rule-Possession*

I discuss whether that a society has a rule simply means that its members believe there is a protected reason to comply with that rule. In its favour, this theory seems capable of explaining nearly everything we know about rule-possession (6.2). It requires, however, that protected reasons, and hence what Raz calls 'exclusionary reasons', take priority over at least some conflicting reasons, regardless of weight. I argue that we have no basis for thinking that

they do (6.3). I do not try to show that the claim is false. But without an argument in its favour, we should not accept that claim as true.

### *Chapter 7 – Acceptance and Counterfactuals*

I argue that exclusionary reasons are part of a group of reasons all of which require actions that can be described in counterfactual terms (7.2). Another member of that group is a reason to act as if you believe something you actually do not. To show how you can bring yourself to satisfy such a reason, I introduce a distinction commonly drawn in the philosophy of action between belief and acceptance (7.3). Acceptance and belief are both propositional attitudes, alike in the way they lead you to act, but neither attitude entails the other. They differ in that acceptance, unlike belief, is shaped by practical reasons; it is under your direct, voluntary control; and it is held relative to certain contexts. I call acceptance in the absence of a corresponding belief ‘make-belief’.

### *Chapter 8 – Acceptance and Evidence Law*

I show that acceptance, as distinct from belief, already has a place in law. I start by explaining some basic features of the law of evidence (8.2). I then explain that judges and juries are often required not to proceed based on what they believe, and these are in substance requirements not to accept what is believed (8.3). They are also often required to presume certain propositions whether or not they believe them, and these are in substance requirements to accept what may not be believed (8.4).

## *Chapter 9 – Acceptance and Rule-Possession*

I propose a new theory of rule-possession. I show that acceptance of a certain kind leads a person to act in the way that having a rule does, and I explain how acceptances of various kinds lead to possession of rules of various strengths and strictnesses (9.2). I then show that rule-possession shares other features of acceptance, including voluntariness, responsiveness to practical reasons, context-dependence (9.3). I conclude by setting out the acceptance theory of rule-possession, listing its advantages, and defending it against some possible objections (9.4).

## *Chapter 10 – Reasons, Possession, and Compliance*

Having answered my primary question – i.e., ‘what is it to have a rule?’ – I conclude by discussing the implications of the answer for some other questions about mandatory rules. I show that instrumental reasons for having rules are typically reasons to make-believe that there is a reason for some action (10.2). I also show that whether you ought to comply with a rule is independent to whether you ought to have a rule, meaning it is possible that you ought to break a rule that you ought to have (10.3).

# 2 Reasons

---

## 2.1 INTRODUCTION

To understand what it is to have a rule, it helps to have in place a basic account of reasons and, relatedly, what it means that something ought to be done. This is mainly because having a rule guides speech and action in much the same way as a belief about what ought to be done does. There are also debates about when having a rule is justified, and when compliance with a rule is justified, and these debates make more sense if we know something about reasons and ‘ought’. My account is based on Joseph Raz’s theory of practical reason, because it is widely accepted in legal philosophy.

The plan for this chapter is as follows. In Section 2, I set out Raz’s account of explanatory reasons. Section 3 is a discussion of normative reasons. I outline Raz’s account of what normative reasons are, state an objection to that account, and then propose a revised account. In Section 4, I look at two ways in which normative and explanatory reasons are related. That section is also where I introduce exclusionary reasons. In Section 5, I discuss conflicts between reasons, and one of the relations between reasons and ‘ought’.

## 2.2 EXPLANATORY REASONS

A reason, whether explanatory or normative, is a fact; it is something of which it can be said ‘it is a fact that ...’.<sup>1</sup> The fact that  $p$  may be a normative reason and also an explanatory reason. The one fact may have these two properties.

What kind of facts are explanatory reasons? An explanatory reason is ‘any fact which figures (non-redundantly) in an explanation’, where an explanation is a correct answer to a why-question.<sup>2</sup> Why do panthers have long tails? Why is there a law against murder? Correct answers to these questions state the reasons why things are the way they are; they explain those things. It is possible, of course, to ask a why-question without using the word ‘why’. We could equally ask: What explains the fact that panthers have long tails? What accounts for there being a law against murder? There may also be alternative explanations for the same object, which are independent, without being in competition with each other.<sup>3</sup>

Just as we can ask for an explanation of something natural, like panthers’ long tails, we can sensibly ask for an explanation for a particular action. An explanatory reason for why someone performed an action is a fact that figures in an explanation of why the person performed it. It is a fact that figures in a correct answer to a question of the form, ‘why did  $A$   $\phi$ ?’ For example, a correct answer to the question ‘why did you take your umbrella?’, might include that you were going outdoors, that you live in England, that the umbrella was placed near the door, and so on.

---

<sup>1</sup> Joseph Raz, *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990) 17-18.

<sup>2</sup> Joseph Raz, ‘Reasons: Explanatory and Normative’ in Constantine Sandis (ed), *New Essays on the Explanation of Action* (Palgrave Macmillan, 2009) 185.

<sup>3</sup> Raz, ‘Explanatory and Normative’ (n 2) 186.

## 2.3 NORMATIVE REASONS

### *Description*

What does it mean that a fact is a normative reason? Raz approaches the question through an account of reason-giving statements and inferences. He describes two basic forms of reason-giving statements:

- (1) There is a reason for  $\mathcal{A}$  to  $\phi$ .
- (2) The fact that  $p$  is a reason for  $\mathcal{A}$  to  $\phi$ .<sup>4</sup>

The question of what it means for a fact to be a normative reason is a question about the meaning of statements of type (2). ‘Statements of the form “ $p$  is a reason for [ $\mathcal{A}$ ] to  $\phi$ ” correspond to an inference of which “ $p$ ” is the premise and “there is a reason for [ $\mathcal{A}$ ] to  $\phi$ ” the conclusion’.<sup>5</sup> The statement is true so long as the inference is sound and valid.<sup>6</sup> This means that a fact is a normative reason for  $\mathcal{A}$  to  $\phi$  if and only if it is stated by a premise of a sound inference the conclusion of which can be stated as ‘there is a reason for  $\mathcal{A}$  to  $\phi$ ’. The conjunction of all the facts stated by such the premises of such an inference is called a ‘complete normative reason’.<sup>7</sup>

---

<sup>4</sup> Raz, *Practical Reason and Norms* (n 1) 29.

<sup>5</sup> Raz, *Practical Reason and Norms* (n 1) 28. See also: Raz, *Practical Reason and Norms* (n 1) 187: ‘We can think of [reasons for an action] as facts statements of which form the premises of a sound inference to the conclusion that, other things being equal, the agent ought to perform the action’.

<sup>6</sup> Joseph Raz, ‘Introduction’ in Joseph Raz (ed), *Practical Reasoning* (OUP 1978) 5; Raz, *Practical Reason and Norms* (n 2) 187.

<sup>7</sup> Raz, *Practical Reason and Norms* (n 1) 24; Raz, ‘Introduction’ (n 6) 5.

### *Dancy's Objection*

According to the definition just given, a normative reason is a fact stated by the premise of a sound, valid inference the conclusion of which is that there is a reason. Jonathan Dancy objects that the premises of some such inferences do not state reasons. Here is Dancy's example:

- (1) Someone in the room has a reason to  $\phi$ .
- (2) There are only three people in the room,  $A$ ,  $B$ , and  $C$ .
- (3) Neither  $B$  nor  $C$  has a reason to  $\phi$ .
- (4) So, there is a reason for  $A$  to  $\phi$ .

The argument is valid and we can assume that it is sound. On the above definition, it follows that (1), (2), and (3) each state a reason for  $A$  to  $\phi$ , and together they state a complete reason for  $A$  to  $\phi$ . But it seems they do not. As Dancy says, '[a]ll we have here is a guarantee that there is a reason, which is not at all the same thing as a specification of the reason that is guaranteed to exist'.<sup>8</sup>

The problem is that sometimes it is valid to infer that there is a reason from evidence that there is a reason, as in Dancy's example. On our definition, that has the unattractive consequence that evidence of a reason for  $A$  to  $\phi$  is itself a reason for  $A$  to  $\phi$ . It would be better to say that a fact is a normative reason for  $A$  to  $\phi$  just if it makes it the case (perhaps in conjunction with other facts) that there is a normative reason for  $A$  to  $\phi$ . In other words, for a fact to be a reason for an action it must be a fact because of which there is such a reason. By similar reasoning, we could say that a complete normative reason is the conjunction of all the facts that (non-redundantly) make it the case that there is a normative reason. These

---

<sup>8</sup> Jonathan Dancy, 'Enticing Reasons' in R Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (OUP 2004) 98. For discussion see John Broome, 'Reasons' (same volume) 36.

definitions are not vulnerable to Dancy's counter-example, because none of (1), (2), or (3) make it the case that there is a reason for  $\mathcal{A}$  to  $\phi$ , nor do they do so together.

### *Reasons and Values*

According to the revised definition, a normative reason is a fact that makes it the case that there is a normative reason. That is not very enlightening on its own. We want to know what kinds of facts make it the case that there is a normative reason. In Raz's theory, the answer starts with the relation between reasons and values. According to Raz, 'practical inferences are governed by the logic of satisfactoriness not that of satisfaction'.<sup>9</sup>

According to the logic of satisfaction it follows from the premisses [of a practical inference] that one ought to perform an action whose performance is *necessary* for the achievement of the goal [or value] the premisses specify. By the logic of satisfactoriness the conclusion specifies an action performance of which *suffices* for the realization of the goal.<sup>10</sup>

That there is a reason for an action implies that its performance realizes some value (or a valid goal or a good).<sup>11</sup> Things also work the other way around. '[O]ne has reason to do whatever is a sufficient condition for the realization of some good'.<sup>12</sup>

Given this rough understanding of the relationship between reasons, actions, and values, any explanation of why there is a reason to  $\phi$  must include the fact that there is a value that  $\phi$ ing realizes. Put another way, a complete normative reason to  $\phi$  must include the fact of

---

<sup>9</sup> Raz, 'Introduction' (n 6) 9. The distinction is due to Anthony Kenny, 'Practical Inference' (1966) 26 Analysis 65.

<sup>10</sup> Raz, 'Introduction' (n 6) 9. It is clear from the context that Raz adopts the 'logic of satisfactoriness'. See, for example: Raz, 'Introduction' (n 6) 11: 'The logic of satisfactoriness is superior to the logic of satisfaction'.

<sup>11</sup> Raz, *Practical Reason and Norms* (n 1) 33.

<sup>12</sup> Raz, *Practical Reason and Norms* (n 1) 182. Raz accepts that there are reasons to take even extreme means to realize trivial values. In one of his examples, he says that if blowing up a house is a means of killing a fly, and killing the fly is a valid goal, then there is a reason to blow up the house. '[I]n so far as killing the fly is concerned there is indeed nothing wrong with blowing up the house. We regard this as absurd only because of the other bad consequences of the action'. Raz, 'Introduction' (n 6) 11.

a value that  $\phi$ ing realizes. Also, if we think of realizing a value as an action trivially sufficient to realize a value, then the fact of a value is a complete normative reason for that value's realization.<sup>13</sup> The fact of a value is called an 'operative reason'.<sup>14</sup> Thus, 'every complete reason includes an operative reason and ... every operative reason is a complete reason for some action or other'.<sup>15</sup>

In addition to an operative reason, a complete reason may include the fact that (or the facts that make it the case that) a certain act-type realizes a value. Such facts are 'auxiliary reasons'. They are 'facts pointing to a way of realizing the goal (or value)' the fact of which is the operative reason.<sup>16</sup> Together, the operative and auxiliary reasons make up a complete reason for the act-type identified by the auxiliary reason'.<sup>17</sup> Here is an example of how the two types of reasons combine:

I want to help him. Lending him £400 will help him. Therefore, I have a reason to lend him £400. The first premise states an operative reason; the second states an (auxiliary) reason. It transmits, as it were, the force of the operative reason to the particular act of lending him £400.<sup>18</sup>

Together, the fact you want to help him, plus the fact lending him £400 will help him, form a complete reason to lend him £400. The same operative reason can combine with other auxiliary reasons to form other complete reasons. For instance, the fact that you want to help him, and the fact that driving him to the bank will help him, form a complete reason to drive

---

<sup>13</sup> Raz, *Practical Reason and Norms* (n 1) 33: 'It seems to me a logical truth ... that every operative reason is a complete reason for some action or other'. On the same page, he gives these examples of 'complete reasons consisting of one operative reason of the actions for which they are reasons': 'If respect for persons is a value then there is a reason for everyone to respect persons. ... If [one] desires x then there is reason for him to promote the realization of his desire, etc'.

<sup>14</sup> Raz, 'Introduction' (n 6) 15. Raz thinks that norms, including rules, are also operative reasons: Raz, *Practical Reason and Norms* (n 1) 34, 79-80.

<sup>15</sup> Raz, *Practical Reason and Norms* (n 1) 33.

<sup>16</sup> Raz, 'Introduction' (n 6) 15.

<sup>17</sup> Raz, *Practical Reason and Norms* (n 1) 34.

<sup>18</sup> Raz, *Practical Reason and Norms* (n 1) 34-35.

him to the bank. Also, a complete reason can include more than one auxiliary reason. Here is another example:

(1) *G* is in my interest; (2) *P* is sufficient for *G*; (3) doing *A* will (or is likely to bring about that *P*; it follows that relative to (1), (2), (3) I ought to do *A*. Here (1) states an operative reason. (2) and (3) state auxiliary reasons. (2) and (3) are by themselves no reason for any action.<sup>19</sup>

In this case, both (1) – the operative reason – and (1), (2) and (3) together are complete reasons.

### *Summary*

In this section I presented Raz's definition of a normative reason, raised Dancy's objection that definition, and offered a revised definition in its place. In the revised definition, explanatory and normative reasons came together, such that a fact is a normative reason if it makes it so there is a normative reason. In the next section, I discuss some of the other possible explanatory roles for normative reasons, in particular how they can explain the actions for which they are reasons.

## 2.4 RELATIONS BETWEEN EXPLANATORY AND NORMATIVE REASONS

### *Acting for a Reason*

When you perform a particular act we can ask why you performed it. One standard type of explanation refers to a reason for which you performed the action. What is it to act for a

---

<sup>19</sup> Raz, 'Introduction' (n 6) 15.

reason? Acting for a reason involves believing there is some normative reason for an action and, as a result, performing that action.<sup>20</sup> There is a great deal of debate about the way in which the belief must explain the action, but that is beyond the scope of my analysis. I will do what is customary and say that it must explain the action in the ‘right way’.<sup>21</sup> Putting things together, I will say that you  $\phi$  for  $p$  if and only if you  $\phi$  because you believe that  $p$  and that  $p$  is a normative reason to  $\phi$ , and the explanatory connection works in the right way. Sometimes, I will say that a person  $\phi$ s ‘for the reason that  $p$ ’ instead of saying he  $\phi$ s ‘for  $p$ ’.<sup>22</sup> According to this definition, when you act for a reason, your belief in a normative reason is an explanatory reason for your action.

There are several things to note about this definition of acting for a reason. First, you can believe there is a reason for an action, where that belief explains why you perform the action, without your being aware of the belief.<sup>23</sup> Second, to act for a certain reason you need to perform the action you believe that reason favours. It is not enough to perform some other action.<sup>24</sup> Suppose you believe it is raining and that it is raining is a reason to take your umbrella. Because you believe that, you place your umbrella near the door. In placing your

---

<sup>20</sup> I do not mean that to act for a reason you must have the concept of a normative reason. You may not be able to say what is a normative reason, or be able to think about a normative reason abstractly. But you will be able to use the concept, in that you will be able to treat facts as normative reasons.

<sup>21</sup> Many of the problem cases involve ‘deviant causal chains’, where a belief that there is a reason  $p$  to  $\phi$  causes someone to  $\phi$ , but the explanation does not work through the person’s rational faculty. In such cases we would not say that the person  $\phi$ -s for  $p$ , but identifying the right kind of explanatory connection has proved extremely difficult. See: Donald Davidson, *Essays on Actions and Events* (Oxford 1980) 79ff; George Wilson, *Intentionality of Human Action* (Stanford University Press 1989) ch 9; Alfred Mele, *The Springs of Action* (Oxford 1992) ch 11; Kieran Setiya, *Reasons without Rationalism* (Princeton University Press 2007) 31-32.

<sup>22</sup> Raz says that ‘x’s reason for  $\phi$ -ing was  $p$ ’ means ‘x believed both that  $p$  and that  $p$  is a reason for him to  $\phi$ , and x  $\phi$ -ed intentionally because of these beliefs of his’. Raz, *Practical Reason and Norms* (n 1) 21. He also says that ‘a person  $\phi$ -s for the reason that  $p$  if, and only if, he  $\phi$ -s because he believes that  $p$  is a reason for him to  $\phi$ ’. Raz, *Practical Reason and Norms* (n 1) 39. Raz cautions that the ‘because’ is problematic; but even so, neither formulation is exactly what I want. The second formulation omits the requirement that the person believes that  $p$ , as well as believing that  $p$  is a reason to  $\phi$ . The first formulation includes that the person  $\phi$ ’s intentionally, which I take to be implied by acting for a reason.

<sup>23</sup> Joseph Raz, *Engaging Reason: On the Theory of Value and Action* (OUP 1999) 231: ‘Action for reasons presupposes an appreciation of the situation we are in, but it does not presuppose self-knowledge.’

<sup>24</sup> Heidi Hurd seems to make this mistake when she claims that deliberating on the basis of a reason may be one of the actions ‘barred’ by a reason not to act for that reason. Heidi Hurd, *Moral Combat: The Dilemma of Legal Perspectivalism* (Cambridge University Press 1999) 77 n 23. See also, Michael Moore, ‘Authority, Law, and Razian Reasons’ (1989) 62 Southern California Law Review 829, 883-885.

umbrella near the door, you may have acted for a reason, but you did not act for the reason to take your umbrella.

Finally, that you act for a reason does not entail that there is that reason. You can act for  $p$  even though it is not the case that  $p$  or it is not the case that  $p$  is a reason for the action you perform. Admittedly this can sound a bit odd. It would be odd to say ‘I took my umbrella for the reason it is raining, but it is not raining’. On this definition, however, it is not incorrect. All I can say is that Raz and others seem to be untroubled by the oddity, and that there are at least sometimes less jarring ways of saying the same thing, such as by stating the intention or purpose with which someone performed an action.<sup>25</sup>

### *Exclusionary Reasons*

Sometimes there is a reason as well as a reason for or against acting for that reason. Reasons for and against acting for a reason ‘positive’ and ‘negative second-order reasons’.<sup>26</sup> Other reasons are called ‘first-order reasons’ (though I will just call them ‘ordinary’ reasons). Negative second-order reasons have a special name: ‘exclusionary reasons’.<sup>27</sup> Exclusionary reasons are important to Raz because they are supposed to always ‘prevail’ over the reasons they exclude. I will say more about this issue in the next section and in Chapter 6. For now, the point to note is that the strong role Raz gives to exclusionary reasons makes it important to know what exactly they are reasons to do.

---

<sup>25</sup> Raz says that ‘[a]cting with an intention or a purpose is acting (as things appear to one) for a reason’. Raz, ‘Explanatory and Normative’ (n 2) 1

<sup>26</sup> Raz, *Practical Reason and Norms* (n 1) 39.

<sup>27</sup> Raz, *Practical Reason and Norms* (n 1) 39.

Suppose  $p$  is a reason to  $\phi$  and  $e$  is a reason not to act for  $p$ . You satisfy  $e$  – meaning, you do as  $e$  is a reason to do – under either of two conditions.<sup>28</sup> First, you satisfy  $e$  if you do not  $\phi$ . That is simply because you do not  $\phi$  for  $p$  if you do not  $\phi$  at all. Second, you satisfy  $e$  if you  $\phi$  not for  $p$ . You might  $\phi$  for another reason or for no reason. Either way, you satisfy the exclusionary reason because the requisite explanatory connection between your belief in  $p$  as a reason to  $\phi$  and your  $\phi$ ing is lacking. More generally, you satisfy both an exclusionary and an excluded reason if you do what the excluded reason is a reason to do but not for that reason.<sup>29</sup>

Here is a version of an example of Raz's. The pleasure in eating seems to be a reason for eating. There might also be a reason not to eat for pleasure. Supposing there are such reasons, the reason not to eat for pleasure is an exclusionary reason that excludes the reason for eating. It is possible to satisfy the exclusionary reason by not eating, in which case the reason for eating goes unsatisfied. In addition, as Raz says, '[o]ne can satisfy both [reasons], for one can eat without eating for pleasure. (Notice that one can have pleasure in eating even when one does not eat for pleasure)'.<sup>30</sup>

I have said what an exclusionary reason *is* a reason to do. Now let me say what an exclusionary reason is *not* a reason to do. An exclusionary reason is not a reason against using the excluded reason in your deliberations about what to do.<sup>31</sup> It is possible to deliberate with a reason without satisfying that reason, so it is possible to deliberate with a reason without

---

<sup>28</sup> As I use the term, to 'satisfy' a reason is, in Raz's terminology, to 'conform' to a reason. Raz, *Practical Reason and Norms* (n 1) 178.

<sup>29</sup> Raz, *Practical Reason and Norms* (n 1) 189.

<sup>30</sup> Joseph Raz, 'Facing Up: A Reply' (1989) 62 *Southern California Law Review* 1153, 1167.

<sup>31</sup> Raz would agree that an exclusionary reason is not (and does not entail) a reason against using the excluded reason in your *theoretical* deliberations. For example, he says that '[s]o long as [the exclusionary reason is complied with], there is little objection to actually engaging in thought about the matter'. Raz, *Practical Reason and Norms* (n 1) 184; see also Joseph Raz, *The Morality of Freedom* (OUP 1984) 39-40; Joseph Raz, *The Authority of Law: Essays on Law and Morality* (OUP 1979) 24-26. Whether Raz would agree that an exclusionary reason is not (and does not entail) a reason against using the excluded reason in your *practical* deliberations is not clear.

acting for that reason. It is also possible to act for a reason without deliberating with it, for example, when you act without deliberating at all. In short, deliberating with a reason is neither necessary nor sufficient to act for it.

Nor is an exclusionary reason a reason against doing just anything because of the excluded reason. It is only a reason against doing what the excluded reason favours for that reason. Suppose that there is a reason to take your umbrella. Taking the umbrella for that reason counts as acting for that reason. Nothing else does. Putting your umbrella by the door does not count as acting for that reason, even if your belief in the reason to take your umbrella explains why you put your umbrella there. Were there an exclusionary reason not to act for the reason to take your umbrella, that reason would not be violated by your putting the umbrella by the door, whatever the explanation for why you did so.

There is one final point. Raz says that there is an exclusionary reason not to act for  $p$  only if it is actually the case that  $p$  and that  $p$  is a reason.<sup>32</sup> In his words, ‘exclusionary reasons exclude valid reasons’.<sup>33</sup> This seems to be a stipulation, so I will take it as given. To be clear, you might still have a reason not to act for  $p$ , even though it is not the case that  $p$  or  $p$  is not a reason; it just would not be an exclusionary reason. The difference matters because of the role exclusionary reasons – but not reasons not to act for an invalid reason – play in conflict situations.

---

<sup>32</sup> Raz, *Practical Reason and Norms* (n 1) 184.

<sup>33</sup> Raz, *Practical Reason and Norms* (n 1) 212. Raz then says: ‘One needs no special reason to exclude bad reasons. Their invalidity itself disqualifies them from being legitimate guides to action.’

## 2.5. CONFLICTS BETWEEN REASONS

A statement of the form ‘ $p$  is a reason for  $\mathcal{A}$  to  $\phi$ ’ is equivalent to a statement of the form ‘ $\mathcal{A}$  ought to  $\phi$  relative to  $p$ ’.<sup>34</sup> A statement of the form ‘there is an undefeated reason for  $\mathcal{A}$  to  $\phi$ ’ is equivalent to a statement of the form ‘ $\mathcal{A}$  ought to  $\phi$  all things considered’.<sup>35</sup> Normally, we want to know what ought to be done all things considered, not just relative to this or that consideration. The important question is therefore when a reason is undefeated. That  $q$  ‘defeats’  $p$  means that  $p$  is a reason for some action but  $p \& q$  is not a reason for that action.<sup>36</sup> Some defeating facts are reasons; others are not. If the defeating fact is not a reason, that fact is said to ‘cancel’ the defeated reason.<sup>37</sup> Raz says little about cancellation, and I will leave it aside. My concern is with reasons that defeat other reasons. When I refer to a reason being ‘defeated’, I mean defeated by another reason. Reasons defeat other reasons in one of two ways, depending on how they conflict.

### *Strict Conflict*

Most conflicts between reasons are ‘strict’ conflicts. Two reasons strictly conflict if and only if the first is a reason to perform an action and the second is a reason not to perform that action.<sup>38</sup> On its own, this definition is quite narrow, but there are two additional principles. Here is the first: ‘If  $p$  is a reason to  $\phi$  and  $q$  is a reason to perform another action,  $[\psi]$ , and it is

---

<sup>34</sup> Raz, ‘Introduction’ (n 6) 13.

<sup>35</sup> Raz, ‘Introduction’ (n 6) 13.

<sup>36</sup> Raz, ‘Introduction’ (n 6) 12.

<sup>37</sup> Raz, *Practical Reason and Norms* (n 1) 27; Raz, ‘Introduction’ (n 6) 12-13.

<sup>38</sup> Raz, *Practical Reason and Norms* (n 1) 25. Here is Raz’s exact formulation: ‘ $p$  strictly conflicts with  $q$  relative to  $x$  and  $\phi$  if, and only if,  $R(\phi)p,x$  and  $R(\sim\phi)q,x$ , i.e. that  $p$  is a reason for  $x$  to  $\phi$  and that  $q$  is a reason to refrain from  $\phi$ -ing’, where ‘ $R(\phi)p,x$ ’ means ‘ $p$  is a reason for an agent  $x$  to  $\phi$ ’.

logically impossible both to  $\phi$  and to  $[\psi]$ , then  $q$  is also a reason to refrain from  $\phi$ -ing and it strictly conflicts with  $p$ .<sup>39</sup> Essentially, the principle says that a reason for one action entails a reason not to perform a logically incompatible action. The second principle is similar and addresses physically incompatible actions: 'If ... it is ... physically impossible both to  $\phi$  and to  $[\psi]$ , then  $q$ , in conjunction with the facts which make it impossible to perform both actions, is a reason to refrain from  $\phi$ -ing'.<sup>40</sup> Together these principles mean that reasons for logically or physically incompatible actions either strictly conflict or are part of reasons that strictly conflict.

If one of two strictly conflicting reasons is 'weightier' or more important than the other reason, then normally it defeats the other reason, and is not defeated by it. The first reason is said to 'override' the second reason.<sup>41</sup> The resolution of strict conflicts also determines what the 'balance of reasons' favours. The balance of  $p$  &  $q$  favours the action that  $p$  favours, so long as  $p$  is not overridden by  $q$ . The overall balance of reasons favours an action, so long as there is a reason for that action which no reason overrides.<sup>42</sup> The important thing to remember is that, when all conflicts are strict conflicts, you ought to do what the overall balance of reasons favours.

### *Partial Conflict*

Not all conflicts are strict conflicts. Exclusionary reasons conflict with the reasons they exclude, according to Raz, but it is possible to satisfy both reasons, so the conflict is not strict.

---

<sup>39</sup> Raz, *Practical Reason and Norms* (n 1) 26.

<sup>40</sup> Raz, *Practical Reason and Norms* (n 1) 26.

<sup>41</sup> Raz, *Practical Reason and Norms* (n 1) 26.

<sup>42</sup> Joseph Raz, 'Reasons for Action, Decision, and Norms' (1975) 84 *Mind* 481, 484.

Nor are all conflicts resolved by weight or strength. Raz says that conflicts between exclusionary and excluded reasons ‘are resolved not by the strength of the competing reasons but by a general principle of practical reasoning which determines that exclusionary reasons always prevail, when in conflict with first-order [i.e., ordinary] reasons’.<sup>43</sup> So, exclusionary and excluded reasons conflict, but not strictly, and exclusionary reasons prevail<sup>44</sup> in such conflicts, but not in virtue of their weight.

When he introduced exclusionary reasons in *Practical Reason and Norms*, Raz did not say exactly how exclusionary and excluded reasons conflict. Nor did he argue for the ‘general principle’ in virtue of which exclusionary reasons take priority. He returned to these issues in the postscript to the second edition of that book. There he said that exclusionary reasons ‘partially conflict’<sup>45</sup> with excluded reasons and that they prevail over exclusionary reasons due to ‘general considerations affecting partial conflicts’.<sup>46</sup> I will wait to discuss partial conflicts and their significance for exclusionary reasons until Chapter 6, when it will be clear why that topic deserves a detailed treatment.

For the moment, it is enough to see how Raz’s ideas about exclusionary reasons fit with his claims about conflicts generally. As I said, what you ought to do all things considered is what there is an undefeated reason to do, which is normally what the overall balance of reasons favours. If a reason is excluded and hence defeated, however, you ought to do what the *unexcluded* balance of reasons favours. These two standards may diverge if the excluded

---

<sup>43</sup> Raz, *Practical Reason and Norms* (n 1) 40:

<sup>44</sup> When I say one reason ‘prevails’ over another reason, I mean the first reason defeats and is not defeated by the second reason. This is consistent with Raz’s usage. Raz also uses a variety of equivalent terms, e.g., ‘win’, ‘triumph’, ‘upper hand’, ‘priority’. See Raz, *Practical Reason and Norms* (n 1) 40, 189, 190; Raz, ‘Facing Up’ (n 30) 1167, 1168.

<sup>45</sup> Raz, *Practical Reason and Norms* (n 1) 189. On partial conflicts generally see: Joseph Raz, ‘Reasons, Requirements, and Practical Conflicts’ in Stephen Korner (ed), *Practical Reason* (Yale University Press 1974). Raz summarizes his discussion of partial conflicts in *Practical Reason and Norms* (n 1) 202 n 4.

<sup>46</sup> Raz, *Practical Reason and Norms* (n 1) 189. See also Raz, ‘Facing Up’ (n 30) 1167-1168.

reason or reasons tipped the overall balance of reasons.<sup>47</sup> On Raz's theory, it is therefore possible, though not certain, that what ought to be done is what the overall balance of reasons opposes.

It may help to consider an example.

Imagine a revolutionary group taking a vow to disregard the government and its laws. ... Their vow is not for them a reason to do the opposite of what the law requires. It is not a reason to break every contract, trespass on all land, steal every property, violate every copyright, etc. It is simply a reason not to perform any action for the reason that the law requires it. One of the revolutionaries considers whether to steal his neighbour's money for revolutionary purposes. He weighs the reasons for so doing (facilitating some revolutionary activities) and against so acting (1) the harm to his neighbour, (2) the law, and (3) the risk to himself if caught. On balance he ought not to steal the money but of the reasons against so doing he should exclude the law because of his vow. It may be that if one excludes the law, the remaining reasons against the theft are overridden by the reason for it. If this is so the vow leads to the conclusion that in this case he ought to act against the balance of reasons.<sup>48</sup>

The reasons against stealing are collectively stronger than the reason for stealing. Normally, that would mean the revolutionary ought not to steal. Yet the revolutionary's vow excludes the law. What he ought to do turns on the balance of the remaining, unexcluded reasons. If facilitating revolutionary activities is a weightier concern than avoiding harm to his neighbour and risk to himself, then the revolutionary ought to steal, even though that goes against the overall balance of reasons.

---

<sup>47</sup> Raz, *Practical Reason and Norms* (n 1) 40: 'One ought not to act on the balance of reasons if the reasons tipping the balance are excluded by an undefeated exclusionary reason'. On the next page, Raz says that '[t]he presence of an exclusionary reason may imply that one ought not to act on the balance of reasons. The exclusionary reason may exclude a reason which would have been overridden anyway, but it may also exclude a reason which would have tipped the balance of reasons ...'. Elsewhere, he writes that, due to an exclusionary reason, a person may face a situation in which, '[o]n one level he ought, on the balance of reasons, to perform an act. On another and superior level he ought to do the opposite' ('Promises and Obligations' in Peter Hacker and Joseph Raz (eds), *Law, Morality and Society: Essays in Honour of H.L.A. Hart* (OUP 1977) 224). To say that an action ought to be performed 'on the balance of reasons' just means that there is a reason for that action that is neither overridden or cancelled, i.e., a 'conclusive reason' (Raz, *Practical Reason and Norms* (n 1) 36; Raz, 'Reasons for Actions' (n 42) 484).

<sup>48</sup> Raz, 'Promises and Obligations' (n 47) 222.

There is one last concept to introduce. Raz calls a ‘protected reason’ a fact that is an ordinary reason for an action and a reason not to act for at least some conflicting reasons.<sup>49</sup> (The ordinary reason is protected from conflicting reasons by the exclusionary reason.) In virtue of its exclusionary character, the protected reason prevails over the reasons it excludes, whatever its weight. For convenience, I will call a ‘fully protected reason’ a fact that is an ordinary reason and an exclusionary reason not to act for *any* conflicting reason, and a ‘partially protected reason’ a fact that is an ordinary reason and an exclusionary reason not to act for *only some* conflicting reasons.

## 2.6 CONCLUSION

The main aim of this chapter was to get clear on some basic topics in practical reason. To that end I formulated definitions of an explanatory reason, normative reason, etc. Another aim was to determine when there is a normative reason, which the discussion of reasons and values helped accomplish. I also wanted to arrive at a rough understanding of the relationship between reasons and ‘ought’, relative and non-relative, and I did so through the discussion of conflicts between reasons. The last goal was to pave the way for a discussion of exclusionary reasons and partial conflicts.

Given that most of this thesis is about rules, it may be useful to have an idea of where it draws on the material in this chapter. At the least that will show that I have not discussed reasons unnecessarily. I refer to reasons more or less constantly in chapters 4, 5, 6, 7, 9, and 10. Throughout, the matters explored in this chapter, including the distinction between

---

<sup>49</sup> Joseph Raz, *The Authority of Law* (OUP 1979) 31.

explanatory and normative reasons, the nature of reasons as facts, and the relationship between reasons and 'ought' are both crucial and assumed. More specifically:

- The material on *values* and the *types of normative reasons* is relevant in Chapter 6 (rule-possession) and Chapter 10 (reasons for possessing rules).
- The material on *acting for a reason* and *exclusionary reasons* is relevant in Chapter 6 (rule-possession) and Chapter 7 (reasons to act 'as if').
- The material on *conflicts between reasons* is relevant in chapters 5 and 6 (rule-possession), Chapter 7 (reasons to act 'as if'), and Chapter 10 (reasons to comply with rules).

Some of the material from this chapter is especially relevant in Chapter 6, where I discuss exclusionary reasons and partial conflicts, but there is no section of this chapter that is relevant only in Chapter 6.

# 3 The External Aspect

---

## 3.1 THE EXTERNAL ASPECT: INTRODUCTION

Having laid the groundwork for this thesis in the last chapter, I can start my discussion of the existence of rules in earnest. I shall begin with what Hart calls the account of the existence of rules ‘we are at first perhaps naturally tempted to give’.<sup>1</sup>

It is that to say that a rule exists means only that a group of people, or most of them, behave ‘as a rule’ i.e. *generally*, in a specified similar way in certain kinds of circumstances. So to say that in England there is a rule that a man must not wear a hat in church or that one must stand up when ‘God Save the Queen’ is played means, on this account of the matter, only that most people generally do these things.<sup>2</sup>

The idea is that there is a rule in a society if and only if the members of that society generally do as that (putative<sup>3</sup>) rule requires. Sometimes, Hart uses different terms to describe actions that are generally performed: ‘pattern of behaviour’, ‘convergent behaviour’, ‘regular uniform behaviour’, ‘standard’ behaviour’, ‘group habit’, etc. I will not try to maintain a consistent usage.

Hart gives qualified support to this idea. ‘Plainly [that most people generally do something] is not enough’ for a rule to exist, he says, ‘even though it conveys part of what is meant’.<sup>4</sup> It is not enough because ‘[m]ere convergence in behaviour between members of a social group may exist (all may regularly drink tea at breakfast or go weekly to the cinema)

---

<sup>1</sup> HLA Hart, *The Concept Law* (2<sup>nd</sup> edn, OUP 1994) 9.

<sup>2</sup> Hart, *Concept of Law* (n 1) 9.

<sup>3</sup> In the rest of this chapter, I leave this clarification implicit.

<sup>4</sup> Hart, *Concept of Law* (n 1) 9.

and yet there may be no rule *requiring* it'.<sup>5</sup> The idea still conveys part of what is meant because '[t]here is certainly one point of similarity between social rules and habits: in both cases the behaviour in question (e.g. baring the head in church) must be general though not necessarily invariable; this means that it is repeated when occasion arises by most of the group'.<sup>6</sup> Thus, for a society to have a rule, it is necessary but insufficient that the members of that society generally do as that rule requires. Hart calls a pattern of compliance with a social rule its *external aspect*.<sup>7</sup> Every social rule, in Hart's view, must have an external aspect.

Schematically, the claim is that, for a society  $S$  to have a rule that  $As$  ought to  $\phi$ , it is necessary but insufficient that  $As$  in  $S$  generally  $\phi$ . Before assessing that claim, I need to make several clarificatory remarks.

*First remark.* Hart's claim is that *social* rules have an external aspect, not that *all* rules have an external aspect. He did not always make the distinction clear<sup>8</sup>, and some early commentators read Hart as making the more general claim, which led them to criticize Hart for failing to notice that many enacted rules are generally broken.<sup>9</sup> However, the overall context shows that Hart intended to draw a connection only between social rules and patterns of behaviour.<sup>10</sup>

---

<sup>5</sup> Hart, *Concept of Law* (n 1) 9. David Shwayder has a very thorough discussion of the differences between rules and generalizations in *The Stratification of Behaviour* (Brill 1971) 243-247. He says: 'Rules "apply", while generalizations "hold". One conforms to or acts in violation of a rule, but not a generalization. Generalizations are true or false; not so rules. ... We may (though we need not always) adopt rules; we establish generalizations. Rules may be rejected; generalizations refuted. Rules may be deliberately changed, even if I cannot change them; not so with generalizations' (246).

<sup>6</sup> Hart, *Concept of Law* (n 1) 55.

<sup>7</sup> Hart, *Concept of Law* (n 1) 56.

<sup>8</sup> For example, in Chapter 1 of *The Concept of Law*, Hart first refers to mandatory rules without qualification, but then on the next page refers to mandatory *social* rules, without marking the shift: *Concept of Law* (n 1) 9, 10.

<sup>9</sup> For example, J Kemp, 'Review of *The Concept of Law*' (1963) 13 *Philosophical Review* 188, 190; M Payne, 'Hart's Concept of a Legal System' (1976) 18 *William and Mary Law Review* 287, 310; R Sherwin, 'Opening Hart's concept of Law' (1986) 20 *Valparaiso University Law Review* 385, 396. For useful discussion of the issue, see: AD Wozzley, 'The Existence of Rules' (1967) 1 *Nous* 63.

<sup>10</sup> See, especially, Hart, *The Concept of Law* (n 1) 255.

*Second remark.* Let us say that you *comply* or *conform* to a rule if and only if that rule requires you to perform some action and you do so. You *violate* or *break* a rule if and only if that rule requires you to perform some action and you do not do so. According to these definitions, if you comply with a rule, you do not break it. However, the converse is not true, for if a rule does not require anything of you, then you neither comply with nor violate the rule, whatever you do. Hart's external aspect is a pattern of *compliance* with a rule, not a pattern of *non-violation* of the rule.

*Third remark.* Hart often describes the external aspect of a rule as 'observable'. For example, he says that the existence of a pattern of behaviour in a group is 'merely a fact about the observable behaviour of most of the group'<sup>11</sup> and that the pattern 'consists in the regular uniform behaviour which an observer could record'.<sup>12</sup> In a related context, he writes that it is possible to 'record the regularities of observable behaviour in which conformity with the rules partly consists'.<sup>13</sup> It is not clear whether Hart is assuming that all behaviour is observable or that social rules only regulate the sub-class of behaviour that is observable. Either way, the claim is questionable. There are social rules against adultery and incest, for example, but it sounds odd to say we could observe people *not* doing these things.<sup>14</sup> Similarly, some mental actions (e.g., decisions) might be regulated by social rules, but mental actions are not observable, at least not under normal conditions. As far as I can tell, nothing crucial in Hart's

---

<sup>11</sup> Hart, *Concept of Law* (n 1) 56.

<sup>12</sup> Hart, *Concept of Law* (n 1) 56. Cf. 'Chess players do not merely have similar habits of moving the Queen in the same way with an external observer, who knew nothing about their attitude to the moves which they make, could record'. Hart, *Concept of Law* (n 1) 56.

<sup>13</sup> Hart, *Concept of Law* (n 1) 89. This statement is made as part of an explanation of the 'external point of view', which is different from the external aspect of a mandatory social rule. Hart is making the point that an observer of a society would be able to identify one aspect of a social rule (though not, as I discuss in the next chapter, every aspect). For discussion, see: Brian Bix, 'HLA Hart and the Hermeneutic Turn in Legal Theory' (1999) 52 SMU Law Review 167, 172-176.

<sup>14</sup> Cf. Cristina Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge University Press 2006) 8: '[N]orms can be either prescriptive or proscriptive: In the latter case, we usually do not observe the proscribed behaviour. As anyone who has lived in a foreign country knows, learning proscriptive norms can be difficult and the learning process slow and fraught with misunderstandings and false steps'.

account depends on whether rule-governed behaviour is observable. To better consider the substance of Hart's claim, I will put the issue aside.

### 3.2 NON-COMPLIANCE, MEMBERSHIP, AND EVASION

Hart never directly says why he thinks a social rule must have an external aspect.<sup>15</sup> Maybe he thinks it is so clearly the case that it does not need argument. Certainly few people seem to disagree with Hart. Nonetheless, I want to challenge that consensus. I argue in this section that a rule may exist in a society even though, generally, its members *do not comply* with that rule. Then, in the next section, I argue more strongly that a rule may exist in a society even though its members generally *break* that rule. These arguments are independent.

#### *Membership*

Sometimes, a society has a rule that does not apply to its own members. It applies only to non-members, to outsiders. Because that rule does not apply to members of that society, it does not require anything of them. Because the rule does not require anything of them, the members of that society do not – indeed, cannot – ever comply with that rule. Hence they do not *generally* comply with that rule. Hence general compliance with a rule is not necessary for that rule to exist in a society. That is the argument. Here is an example. Non-Catholics must refrain from receiving communion. That rule applies only to non-Catholics, so it is impossible for a Catholic to comply with it. Yet it is a rule that Catholics have. Catholics would tell non-

---

<sup>15</sup> One explanation suggests itself. In Chapter 4 of *The Concept of Law*, Hart's aim is partly to show the inadequacy of a theory that takes law to be the coercive orders of a person or persons habitually obeyed. Hart begins by explaining what a habit of obedience is, shows that law is continuous even absent a habit of obedience, and then, in the gap creates, develops his theory of social rules. Perhaps Hart did not think he needed to justify giving habit a role in his theory, since it was the centrepiece of the theory he was seeking to displace.

Catholics not to take communion, for instance. So it would seem that Catholics have a rule that they do not generally comply with.<sup>16</sup>

Although there is no pattern of compliance with the communion rule in the society in which it exists, there *is* a pattern of compliance with that rule among those to whom it applies. That is to say, non-Catholics generally do not take communion. Would it not therefore be better to say that a society has a rule only if there is a pattern of compliance with that rule among those to whom it applies, whether or not they belong to the society in which that rule exists? No, because societies often have rules that apply primarily to non-members even though non-members do not generally, if ever, comply with these rules. Catholics have many rules, regarding procreation in particular, that are universally applicable but that non-Catholics do not generally observe.<sup>17</sup>

Another response might be that I have misunderstood Hart, or been uncharitable to him. Surely, you might say, Hart's concern was, or should be understood to have been, rules that exist in a society *and* that apply to its members. The scope of his claim should be understood to encompass only rules of this sort. I have no particular interest in criticizing Hart. If he is best read as making a more limited claim than it at first appears, one which is consistent with my argument thus far, then so much the better. The communion rule could then be understood as a way of showing why this limited claim cannot be generalized. But it is important to realize that, according to this reading of Hart, a society may have a rule with

---

<sup>16</sup> Might the rule govern the *provision* of communion rather than its *reception*? It seems unlikely, because Catholics say that non-Catholics should not take communion, not merely that priests should not give communion to non-Catholics.

<sup>17</sup> Cf. Sanford Levinson, 'Who is a Jew(ish) Justice?' (1989) 10 *Cardozo Law Review* 2359, 2369: "According to classical Judaism, *all* persons, and not only Jews, are under a Noachite commandment not to worship idols; such behaviour is ultimately punishable by death'. Quoted by Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (OUP 1991) 119 n 9. Herman Melville complained about much the same thing when he visited a Polynesian island community: 'I was many times called to order, if I may use the phrase, when I could not for the life of me conjecture what particular offence I had committed'. Herman Melville, *Typee: A Peep at Polynesian Life* (first published 1846, Northwestern University Press 1968) 91.

which its members never comply. (The example of the Ik, later in this section, and the examples in section 3 are still effective against this limited claim.)

Other responses might take issue with my choice of example. Perhaps you think the communion rule is better thought of as a deliberately created rule than as a social rule. Or perhaps the content of the rule is other than I have supposed.<sup>18</sup> I do not want to let anything turn on these issues, so think of the wide variety of rules that apply only to outsiders. They include admission rules, initiation rites, hazing rituals, and requirements of guests and visitors<sup>19</sup>. Surely some of these rules are social rules. So long as even *one* of them is a social rule, then not *every* social rule depends on a pattern of compliance, which is enough for my purposes.

Here is a more detailed example, this time of a membership rule. The example originates with Marek Kaminski, who was a political prisoner in Poland in the 1980s. Kaminski was also a social scientist, and he used his imprisonment as an opportunity to study the rigid caste system that existed in the prison population. The dominant caste was called the ‘grypsmen’. Grypsmen stood apart from other prisoners. They lived by their own set of strict rules, spoke a secret language, and avoided contact with members of lower castes except to exploit them.<sup>20</sup> ‘The norms of this group’, Kaminski said, were ‘the core elements of the prison subculture.’<sup>21</sup> The grypsmen’s dominance of prison life meant that ‘[f]ew inmates entering a prison cell would not like to join this caste’.<sup>22</sup> Yet becoming a grypsman was not

---

<sup>18</sup> For example, it could be that the rule prohibits taking communion unless one is in a state of grace, a condition that non-Catholics do not meet, but that some Catholics do not meet either. Thought of this way, the rule *does* apply to Catholics, so it *is* possible for there to be a pattern of compliance with the rule among Catholics.

<sup>19</sup> Raz writes: ‘Hart’s analysis applies primarily to rules addressed to all the members of the group in which they exist. Many rules are addressed to a particular sub-class of the group in which they exist .... Other rules apply only to people who are *not* members of the group in which the rules exist (rules as to the behaviour of foreigners visiting the country, etc.)’. Joseph Raz, *The Concept of a Legal System* (2<sup>nd</sup> edn, OUP 1980) 148.

<sup>20</sup> Marek Kaminski, *Games Prisoners Play* (Princeton University Press 2004) 34.

<sup>21</sup> Kaminski, *Games Prisoners Play* (n 20) 34.

<sup>22</sup> Kaminski, *Games Prisoners Play* (n 20) 16.

easy. New inmates – casteless ‘rookies’ – were forced to go through a harsh initiation process, the rules of which were passed down from one generation of grypsmen to the next. One step in that initiation process, called ‘baptism’, tested a rookie’s willingness to endure pain.<sup>23</sup> I will call the rule that required rookies to undergo baptism to join the grypsmen the ‘baptism rule’.

Two points matter for my purposes. First, the baptism rule only applied to rookies. Since no prisoner was both a rookie and a grypsman, the rule did not apply to grypsmen. So, no grypsmen complied with the baptism rule.<sup>24</sup> Second, the baptism rule was a rule that existed in grypsmen society. I say that for several reasons: the grypsmen had many other rules that were theirs alone (e.g., there were complex rules of grypsmen argot); the baptism rule restricted membership to the grypsmen and normally we say that a membership rule is a rule *of* the group whose membership it regulates; the grypsmen taught the baptism rule, ensuring its continuation; and it was the grypsmen who monitored and enforced the baptism rule. Putting these two points together, we can say that a rule existed in grypsmen society with which they did not generally comply.

I should be clear that my claim in this section is merely that the existence of a rule in a society is compatible with general *non-compliance* with that rule among the members of that society. I am not yet claiming that the existence of a social rule is compatible with its general *violation*. The grypsmen, for example, do not comply with the baptism rule, but they do not violate it either. More generally, those to whom a rule does not apply neither comply with nor

---

<sup>23</sup> Kaminski describes baptism thus: ‘Deep into the night, in the light of torches made up butter and sheets, rookies are blindfolded and spread on the stools. The *butcher*, inmate-executioner, prepares a special wet towel that is supposed to “break the rookie’s bones” and yet not leave external signs of beating. Surrounded by a circle of blood-thirsty half-naked inmates, the rookie awaits mortal blows. As the slaughter is about to begin, the butcher offers him an option out of the ceremony in exchange for forgoing grypsmen membership. Those who accept, frightened by the performance, immediately get cursed and beaten by the butcher. ... Those who resist the temptation to opt out become eligible for the next stage of initiation. The actual blows paid by the butcher are symbolic. In fact, baptism is a harmless spectacle.’ Kaminski, *Games Prisoners Play* (n 20) 42-43.

<sup>24</sup> Could you say, every prisoner was a rookie before he became a grypsman, so every prisoner who is a grypsman must have complied with the baptism rule at some point? It depends on when the baptism rule came into existence, but even so, this is not an objection. We are interested in whether the grypsmen generally complied with the rule *while they are grypsmen* – while they were part of the society in question. Likewise, we do not generally decide whether Catholics have a rule by taking into account the behaviour of Catholics before they converted.

violate that rule. Let me also be clear that my concern is only with the necessary conditions for the existence of a rule in a society. Although I deny that there must be a pattern of compliance with *every* social rule, I do not deny that there is in fact a pattern of compliance with *most* social rules.

### *Gifts and Evasion*

The communion rule and the baptism rule do not apply to the members of the societies that possess them, which is why the members of those societies do not comply with those rules. There may also be rules that happen not to apply on any occasion to a member of the society that has the rule. I have in mind rules that require people to do things they do not wish to do, but that only require them in certain circumstances. Unsurprisingly, people will try to avoid those circumstances. If they are always successful, the rule will never require anything, so no one will ever comply with it. Allowing that all this is possible, we must allow that a rule can exist in a society absent general compliance with that rule among the members of that society.

Consider the Ik of northeastern Uganda. Traditionally hunter-gatherers, the Ik were displaced from their lands in the 1950s, at which time they began to slowly starve. In his book, *The Mountain People*, Colin Turnbull reports the lengths to which the Ik went to survive.<sup>25</sup> One strategy was to try to ‘help’ someone with the intention of thereby obligating him or her to return the aid at some later, more advantageous date. Rules of reciprocity became ‘weapons,

---

<sup>25</sup> Colin Turnbull, *The Mountain People* (Jonathan Cape 1972). Turnbull’s work is controversial, partly because he was so negative about the Ik. Chapter headings include ‘The Loveless People’ and ‘The End of Goodness’, and the acknowledgments section begins ‘For the Ik, whom I learned not to hate’. There are also some alleged methodological and linguistic errors. For criticism, see, Bernd Heine, ‘The Mountain People: Some Notes on the Ik of North-Eastern Uganda’ (1985) 55 *Africa* 3. I only require that it is possible that there is a society like the one Turnbull purports to describe. It does not matter, for my purposes, whether Icen society is really like Turnbull says it is. If you wish, you can think of it as a detailed, realistic hypothetical.

sharp and aggressive'.<sup>26</sup> However, as Turnbull notes, 'the purpose for which the gift is designed can be thwarted by the non-acceptance of it'<sup>27</sup>, leading the Ik to expend considerable effort to avoid receiving gifts, even going so far as to fix leaking roofs at night to forestall offers of help.<sup>28</sup>

Suppose the Icier rule of reciprocity requires a person to reciprocate if he or she has received a gift. An Ik would comply with that rule if, and only if, he or she receives a gift and then reciprocates. On Turnbull's account, every Ik is trying as hard as possible to evade the rule's application by frustrating others' attempts at gift giving, and every indication is that the Ik were quite skilled at evading the rule.<sup>29</sup> Perhaps they were skilled enough that none of them received gifts. It certainly seems possible. But then it is possible that no Ik ever complied with the rule of reciprocity. (No Ik would have violated the rule either.) There would be a rule in Icier society absent general compliance with that rule among the Ik.

Why do I say that the Ik would still have a rule of reciprocity even if they always evaded its application? Because the existence of that rule would be part of the best explanation of the Ik's evasive behaviour. As Cristina Bicchieri says, '[t]here are ... behaviours that can be explained only by the existence of norms, even if the behaviour prescribed by the norm in question is never observed'.<sup>30</sup> She applies that lesson to the Ik:

Much of the Ik's behaviour can be explained as a successful attempt at eluding existing reciprocity norms. ... Their practices demonstrate that it is not necessary to observe compliance to argue that a norm exists and affects behaviour. ... Nobody is violating the norm, but everybody is trying to avoid situations where they would have to follow it.<sup>31</sup>

---

<sup>26</sup> Turnbull, *The Mountain People* (n 25) 143.

<sup>27</sup> Turnbull, *The Mountain People* (n 25) 143.

<sup>28</sup> Bicchieri, *The Grammar of Society* (14) 9.

<sup>29</sup> Cristina Bicchieri, 'The Fragility of Fairness: An Experimental Investigation of the Conditional Status of Pro-Social Norms' (2008) 18 *Philosophical Issues* 230, 242 (introducing the term 'norm evasion').

<sup>30</sup> Bicchieri, *The Grammar of Society* (14) 9.

<sup>31</sup> Bicchieri, *The Grammar of Society* (14) 9.

We infer the existence of a rule in part because we witness behaviour that makes the most sense if there is such a rule. Normally, that behaviour is compliance with the rule. But that is not essential. It could be evasive behaviour instead, as in the case of the Ik.

I have been relying on an assumption about the kind of rule the Ik have. Let us say that a ‘narrow-scope’ rule requires you to perform an action but only in certain circumstances, whereas a ‘wide-scope’ version of that rule requires you to either perform that action or avoid those circumstances.<sup>32</sup> The distinction makes a difference to when you comply with a rule. Take a rule that requires you to  $\phi$  in circumstances  $C$ . If the rule has a narrow scope, and you avoid  $C$ , then the rule does not require anything, and you neither violate nor comply with that rule, whether or not you  $\phi$ . In contrast, if the rule has a wide scope, and you avoid  $C$ , then you comply with that rule, whether or not you  $\phi$ . In general, you *comply* with a wide-scope rule when you would merely *not violate* a narrow-scope version of that rule.

By assuming that an Ik neither violates nor complies with the rule of reciprocity if he or she does not receive a gift, I have assumed that the rule has a narrow scope. Were the rule to have a wide scope instead, an Ik would comply with that rule if he or she either does not receive a gift or receives a gift and reciprocates. So, were the rule a wide-scope rule, and were it true that the Ik never receive gifts – as I have supposed it is – then all Ik would comply with the rule of reciprocity on all occasions, which, of course, would mean that the Ik generally comply with the rule of reciprocity.

You might then think I have been too hasty. True, if the Icier rule of reciprocity has a narrow scope, it might cause problems for Hart; but if the rule has a wide scope, then not. We do not have any reason for thinking the rule has a narrow or wide scope. So, we do not

---

<sup>32</sup> Cf. John Broome, ‘Normative Requirements’ (1999) 12 Ratio 398.

have any reason for treating the rule as a counterexample to Hart's analysis. That would be one objection to my argument.

This objection gets things backwards, however. If Icier society is the way I have assumed it is – if no Ik receives a gift or reciprocates – then Hart needs the rule of reciprocity to have a wide scope. It is not enough that the rule *could* have a wide scope; it *has* to have a wide scope. Agnosticism about the rule's scope works against Hart, not for him. Hart cannot allow even the possibility that the Ik's rule has a narrow scope, because that concedes the possibility that a society has a rule absent general compliance with that rule among the members of that society – and that possibility is all I am arguing for. Roughly the same reply would be appropriate were the objection to be that I have simply assumed, rather than shown, that the Ik always successfully evade the rule's application. I made that assumption for the sake of argument, to demonstrate that it is possible that the Ik never reciprocate even though, based on their evasive behaviour, we would still say that a rule of reciprocity exists in their society.

### 3.3 VIOLATIONS

In the examples in the last section, a rule exists in a society, but its members do not comply with it, the reason being that the rule does not require anything of them. I now want to briefly argue for the same conclusion but this time on the basis that sometimes a society has a rule even though its members generally break the rule.

I will start with what I think is a flawed example. Rolf Sartorius is one of the few scholars to deny that rules need an external aspect. His argument for that position is very brief. It consists of a single example.

[C]ontrary to what seems to be a virtually universal assumption among philosophers, it makes perfect sense to speak of a social rule as existing in a community even though the external aspect of its existence ... is absent. I take it that contemporary American sexual morality contains a prohibition on adultery. I also think that it might be true that the majority of married individuals commit adultery at least once in their lifetimes. The point is that it could be true without changing the fact of the matter concerning the existence of a moral prohibition on such behaviour.<sup>33</sup>

There is a rule in America against committing adultery, Sartorius says, and either that is or would be the case even though or even if most married Americans have committed adultery and thus have violated the rule at least once. There are other examples along the same lines. There is a rule that requires you to spell out single-digit whole numbers (with some exceptions), but most English speakers have, on occasion, broken the rule.

Sartorius' example shows that a society can have a rule that most of its members occasionally violate. Sartorius thinks it also shows that Hart has made a mistake. Hart's claim is that a society can have a rule only if there is general compliance with the rule. So, Sartorius must be assuming that 'general compliance' means something like 'invariable compliance by most members'. That is not a safe assumption, however, for two reasons. Sometimes we say that people 'generally' act in some way and mean simply that most people act in that way on most (rather than all) occasions. In that less restrictive sense, married people are generally faithful because most of them are faithful most of the time.<sup>34</sup> More importantly, invariable compliance by most members does not seem to be what Hart has in mind. He says that what is necessary for the existence of a rule in a society is that 'most' people 'generally' do as a rule requires<sup>35</sup>; that such conduct must be 'repeated when occasion arises by most of the group'<sup>36</sup>;

---

<sup>33</sup> Rolf Sartorius, 'Positivism and the Foundations of Legal Authority' in 'Positivism and the Foundations of Legal Authority' in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy: The Influence of HLA Hart* (OUP 1987) 51. See also: Geoffrey Warnock, *The Object of Morality* (Methuen & Co Ltd 1971) 47-48.

<sup>34</sup> The point is made by John Finnis in his 'Comment' on Sartorius' paper in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy: The Influence of HLA Hart* (OUP 1987): 'I think [Sartorius'] argument about adultery in contemporary American sexual mores is weak; the fact that most people commit adultery at least once would not nearly show that the rule had no "external aspect"; the proper question (relative to this so-called external aspect) is whether most people *usually* abstain from adultery when occasion offers' (67 n 53). See also the 'Comment' by Hanina Ben-Menachem, same volume.

<sup>35</sup> Hart, *Concept of Law* (n 1) 9.

or that there is a pattern of conduct that is ‘regularly followed by most members of the group’.<sup>37</sup> It is probably more accurate, and certainly more charitable, to read Hart as favouring a less restrictive requirement than Sartorius suggests he does, in which case Sartorius’ counterexample misses its mark.

A cleaner counterexample would show that a rule may exist even though most people violate it on most occasions. Take the rule in our society that you are to shake hands with someone on first meeting them. Most people do so on most occasions, but not always. In particular, people tend not to shake hands when there is a good chance of thereby catching some serious disease. Shaking hands is less common in hospitals than in similar institutions, for instance. We might say that safety concerns ‘override’ the rule – that is to say, they justify the rule’s violation.<sup>38</sup> Now imagine there arrives a deadly plague, communicated by skin contact. Almost everyone stops shaking hands. There is no longer general compliance with shaking hands rule; indeed, most people violate this rule on most occasions.<sup>39</sup> They do so because the rule is overridden on most occasions, for most people. Safety is simply more important than behaving politely. Nonetheless, we know the rule still exists because those who are susceptible to the disease apologize for their impoliteness, those who are immune keep shaking hands, and because everything goes back to normal once the plague passes.<sup>40</sup>

This example is not meant to be conclusive, but it points to a general problem for Hart. As I explain in more detail in Chapter 5, some social rules are weaker, or more easily ‘overridden’, than others. Naturally, people tend not to comply with a rule when it is

---

<sup>36</sup> Hart, *Concept of Law* (n 1) 55.

<sup>37</sup> Hart, *Concept of Law* (n 1) 255. This passage appears in the postscript, which was published after Sartorius’ article.

<sup>38</sup> I discuss override in Section 5.2.

<sup>39</sup> Could it be that the rule includes an exception for safety? Yes, but so long as it *possible* that the rule does not include an exception for susceptible persons, it is *possible* that there is a rule that is not complied with on most occasions.

<sup>40</sup> Warnock has something similar in mind. He thinks it is possible that there is a rule that most people violate. The example he gives is of a rule that has fallen out of use, but is still acknowledged. But he cautions that ‘this would scarcely look like good sense if the putative rule had *never* been complied with, and also that, if general neglect were sufficiently protracted, it would be natural to regard the rule as, so to speak, gradually fading away.’ Warnock, *The Object of Morality* (n 33) 48

overridden. And this has an important consequence. For it to be plausible that *every* social rule is generally complied with, it must also be plausible that *no* social rule is *usually* overridden. But that is not plausible. There are so many trivial social rules – of etiquette, grammar, manners, etc. – that it seems much more reasonable to think that some social rules are overridden on most occasions. At least it is a possibility, as my example is meant to show.<sup>41</sup>

### 3.4 CONCLUSION

In this chapter I began my analysis of the existence of rules. I began with the view that Hart started his discussion with, namely, that a pattern of compliance is necessary and sufficient for the existence of a rule. Hart rejected that idea, in part, and adopted it, in part. I rejected it entirely. First, I argued that a rule may exist in a society absent a pattern of compliance with the rule because not all rules apply to the members of the society to which they belong. Next, I argued for the same conclusion on the ground that some rules are generally broken by the members of the society to which they belong. My conclusion is that the members of a society need not generally comply with a rule for it to be a rule of their society. That does not mean that I deny that there is general compliance with most societal rules.

---

<sup>41</sup> You might wonder if Hart could modify his claim to take into account this possibility. Perhaps he could claim something like, for a society to have a rule, it must be generally complied with on occasions when it is not overridden. I am willing to concede that, if this claim were further modified, it might be difficult or impossible to find a counterexample to it. Remember, though, all along we have lacked a positive argument for thinking that the existence of a social rule depends on compliance (to any degree) with that rule – and this modified claim would be no different.

# 4 The Internal Aspect

---

## 4.1 THE INTERNAL ASPECT: INTRODUCTION

In the last chapter I explained that Hart thinks that a pattern of compliance with a rule in a society is necessary but insufficient for the existence of that rule in that society. I claimed, more strongly, that this pattern of compliance is unnecessary for the existence of that rule. Either way, though, something is missing. Something else is needed for a society to have a rule. This chapter is about what Hart thinks that missing element is.

Hart first identifies what needs explaining. When there is merely a pattern of conduct, Hart says, '[d]eviations from the regular course need not be a matter for any form of criticism'<sup>1</sup>, but 'where there is ... a rule ... deviations are generally regarded as lapses or faults open to criticism, and threatened deviations meet with pressure for conformity'.<sup>2</sup> Moreover, '[c]riticism for deviation [from a rule] is regarded as legitimate or justified ... as are demands for compliance with the standard when deviation is threatened'.<sup>3</sup> These demands, criticisms, and acknowledgements of legitimacy are expressed using normative vocabulary, including 'ought', 'must', 'should', 'right, and 'wrong'.<sup>4</sup>

Hart thinks that these demands, criticisms, and acknowledgements, as well as the use of normative language, all display or manifest a certain 'critical reflective attitude' to the

---

<sup>1</sup> HLA Hart, *The Concept Law* (2<sup>nd</sup> edn, OUP 1994) 55.

<sup>2</sup> Hart, *Concept of Law* (n 1) 55.

<sup>3</sup> Hart, *Concept of Law* (n 1) 55-56.

<sup>4</sup> Hart, *Concept of Law* (n 1) 57. See also: HLA Hart, 'Scandinavian Realism' (1959) 17 *Cambridge Law Journal* 233, 238.

pattern of conduct as a 'standard'. This attitude is the missing piece that, together with a pattern of conduct, entails the existence of a social rule.

What is necessary [for there to be a social rule] is that there should be a critical reflective attitude to certain patterns of behaviour as a common standard, and that this should display itself in criticism (including self-criticism), demands for conformity, and in acknowledgements that such criticism and demands are justified, all of which find their characteristic expression in the normative terminology of 'ought', 'must', and 'should', 'right', and 'wrong'.<sup>5</sup>

Later, in the Postscript to *The Concept of Law*, Hart says that the critical reflective attitude 'consists in the standing disposition of individuals to take ... patterns of conduct both as guides to their own future conduct and as standards of criticism'<sup>6</sup>. The possession of the critical reflective attitude by the members of a society Hart names the 'internal aspect' of a social rule, to go along with the pattern of conduct that is its external aspect.

Hart illustrates the combination of the external and internal aspects of a rule with an example from chess. Chess players regularly move the Queen in a certain way (the rule's external aspect). But they also 'regard [moving the Queen in that way] as a standard for all who play the game'<sup>7</sup> (the rule's internal aspect). 'Each [player] "has views" about the propriety of all moving the Queen in that way'.<sup>8</sup> Those views are 'manifested' in criticisms of others when they do not move the Queen in the relevant way and demands for conformity when deviation is threatened, as well as 'in the acknowledgement of the legitimacy of such criticism and demands when received by others'.<sup>9</sup> Moreover, for the expression of these criticisms and demands a range of normative language is used, e.g., "I (You) ought not to have moved the Queen like that", "I (You) must do that", "That is right", "That is wrong".<sup>10</sup> Together, the

---

<sup>5</sup> Hart, *Concept of Law* (n 1) 56.

<sup>6</sup> Hart, *Concept of Law* (n 1) 255.

<sup>7</sup> Hart, *Concept of Law* (n 1) 57.

<sup>8</sup> Hart, *Concept of Law* (n 1) 57.

<sup>9</sup> Hart, *Concept of Law* (n 1) 57.

<sup>10</sup> Hart, *Concept of Law* (n 1) 57.

pattern of conduct and the players' attitudes mean that there is a rule that the Queen may be moved in a certain way.

In the next section I provide a more formal statement of Hart's account of the existence conditions of social rules. Then I discuss some criticisms of Hart's account. For now I need to make three clarificatory remarks.

*First remark.* Hart never says exactly what it is to possess a critical reflective attitude to a pattern of conduct as a standard. He is clearest about what the attitude is *not*, namely, a feeling.

The internal aspect of rules is often misrepresented as a mere matter of 'feelings' in contrast to externally observable physical behaviour. No doubt ... individuals may often have psychological experiences analogous to those of restriction or compulsion. When they say they 'feel bound' to behave in certain ways they may indeed refer to these experiences. But such feelings are neither necessary nor sufficient for the existence of 'binding' rules.<sup>11</sup>

(The context makes plain that the reference to 'binding' rules' in the last sentence is a reference to social rules.) It is not clear exactly why Hart is adamant that the attitude is not a feeling, though he may think a feeling is not 'critical' or 'reflective' enough to account for the kinds of demands, criticisms, and other statements people make when there is a rule.<sup>12</sup>

*Second remark.* Hart's choice to use 'standard' in the phrase 'take a pattern of conduct as a standard' is somewhat puzzling. Plausibly all rules are standards but, and this is the important point, not all standards are rules.<sup>13</sup> Why, then, would we say that there is a *rule*

---

<sup>11</sup> Hart, *Concept of Law* (n 1) 57; see also 87-88.

<sup>12</sup> Hart may also wish to distinguish his view of rules from the views of Alf Ross and other Scandian realists, according to which the validity of a rule depends on a 'feeling of compulsion'. See: Hart, 'Scandinavian Realism' (n 4) 236-238.

<sup>13</sup> At times Hart seems to suggest that standards and rules are separate species of the same genus (norm, presumably). At other times he seems to think that rules are a species of standard. See Hart, *Concept of Law* (n 1) 168. Neil MacCormick claims that the second view is clearly the correct one: Neil MacCormick, *HLA Hart* (2<sup>nd</sup> edn, Stanford University Press 2008) 56. Whatever Hart's exact view of rules and standards, it is clearly *not* the view taken by Henry Hart and Albert Sacks, according to which rules require for their application only a determination of fact, whereas standards require a 'qualitative' determination: Henry M Hart, Jr and Albert M Sacks, *The Legal Process: Basic Problems in the Making and Application of Law* (William N Eskridge, Jr and Philip P Frickey eds, Foundation Press, 1994) 138-141.

provided that enough people take a pattern of conduct as a standard? If we were going to say anything at all, wouldn't we say there is a *standard* and leave it open whether the standard is a rule? Put another way, why does Hart think that he has identified the existence conditions of a social rule rather than of a social standard?

It seems likely that Hart did not choose the word standard over rule for substantive reasons. I say that, first, because it would help explain why Hart claims to have identified a condition for the existence of a social rule instead of a social standard. Second, Hart uses 'standard' and 'rule' in similar ways without distinguishing them.<sup>14</sup> Third, completing the phrase 'take a pattern of conduct as a ...' with 'standard' instead of 'rule' is less likely to lead to confusion between descriptive and normative rules. As Margaret Gilbert explains, 'Hart's discussion of the meaning of the common phrase "they do it *as a rule*" (which does *not* imply that they regard something as a standard), may have led him to prefer the term "standard here".<sup>15</sup> It seems safe to assume that 'rule' could be substituted for 'standard' without distorting Hart's meaning. To keep the number of terms to a minimum I will not use 'standard' except in quotations.

*Third remark.* The internal aspect of a social rule is the possession of a certain critical reflective attitude by the members of the relevant society. But must *all* members of a society possess that attitude for there to be a social rule, or some proportion of them, or will just one person do? Must certain members possess the attitude or does everyone count equally?<sup>16</sup> Is it

---

<sup>14</sup> For example, Hart sometimes refers to the acceptance of a 'rule' and sometimes to the acceptance of a 'standard' without flagging the difference: 'Scandinavian Realism' (n 4) 238; *Concept of Law* (n 1) 57-58, 116-117, 168-169. Confusingly, he sometimes refers to the acceptance of rules *as* standards. See, e.g., *Concept of Law* (n 1) 117.

<sup>15</sup> Margaret Gilbert, 'Social Rules: Some Problems for Hart's Account, and an Alternative Proposal' (1999) 18 *Law and Philosophy* 141, 146 n 22.

<sup>16</sup> Michael Bayles suggests that only a small minority of judges or other officials need to have the appropriate attitude for a rule of recognition to exist, provided that they are sufficiently influential. '[I]t might be argued that most of the judiciary in Nazi Germany did not take an internal point of view toward a rule of recognition declaring the Fuhrer's orders to be valid law. Instead, they complied with it from fear of reprisal by the Nazi Party if they did not. ... Yet ... some members of the Nazi Party have to accept the rule of recognition to be motivated to enforce it on judges. What is required is that those who do take an internal point of view have the social power to ensure that officials will follow the approved secondary rules.' Michael Bayles, *HLA Hart's Legal Philosophy: An Examination* (Kluwer 1992) 78.

important that some of those to whom the rule applies possess the attitude?<sup>17</sup> For how long must the attitude be possessed and with what degree of constancy? These questions do not much interest Hart. He says:

How many of the group must in these various ways treat the regular mode of behaviour as a standard of criticism, and how often and for how long they must do so to warrant the statement that the group has a rule, are not definite matters; they need not worry us more than the question as the number of hairs a many may have and still be bald. We need only remember that the statement that a group has a certain rule is compatible with the existence of a minority who not only break the rule but refuse to look upon it as a standard either for themselves.<sup>18</sup>

Hart is willing to say that not *all* members of a society must possess the critical reflective attitude for there to be a social rule. Everything else he leaves open.<sup>19</sup> I will say that ‘enough’ or a ‘sufficient number’ of members of a society must possess the critical reflective attitude for a society to have a rule.<sup>20</sup>

## 4.2 RAZ’S SUMMARY

According to Hart, a rule exists in a society just if there is a pattern of conduct in that society and enough of the society’s members possess a certain attitude, and Hart says that this attitude is manifested in relevant demands, criticisms, and acknowledgements. Many legal

---

<sup>17</sup> No, according to Nick Barber. ‘One section of a society may believe a rule is a standard for another group, and, by enforcing the rule, cajole that group into following the rule even though no members of the subject group believe in its validity. Social rules regulating the conduct of minorities in deeply prejudiced societies may sometimes be of this type.’ NW Barber, *The Constitutional State* (OUP 2010) 62.

<sup>18</sup> Hart, *Concept of Law* (n 1) 56.

<sup>19</sup> That is Hart’s explicit position, but some of his comments muddy the waters. For instance, in *Concept of Law* (n 1) 56, Hart says that ‘if a social rule is to exist *some at least* must look upon the behaviour in question as a general standard to be followed by the group as a whole’ (emphasis added), but on the same page he says that for there to be a social rule, criticism of deviations from the standard and demands for compliance must be regarded as legitimate ‘except by a *minority* of hardened offenders’ (emphasis added). Stephen Perry, describing Hart’s views, writes that ‘[t]he internal point of view is the point of view of those in the group, *consisting of at least a majority*, who accept that the rule is binding upon all members of the group in the manner suggested by the rule’s internal aspect’: ‘Hart on Social Rules and the Foundations of Law: Liberating the Internal Point of View’ (2006) 75 *Fordham L Rev* 1171, 1179 (emphasis added).

<sup>20</sup> Cf. Geoffrey Warnock, *The Object of Morality* (Methuen & Co Ltd 1971) 47: ‘[W]hat could identify ... behaviour as rule-breaking? It seems the only possible answer is: the attitudes of people towards that behaviour. But then, of what people? And here one cannot, I think, do better than say: of enough people, or perhaps, in some cases, of the right people’.

philosophers also read Hart as saying that these kinds of statements are conclusive of the attitude's possession. (Or, at any rate, they think Hart can fairly be presented that way.) According to this reading, it makes sense to think that Hart's claim is that a rule exists in a society just if there is a pattern of conduct and enough of the society's members make certain demands, criticisms, and acknowledgements. This is how Raz seems to understand Hart, for example.

Hart's theory is known as the 'practice theory'. Here is Raz's summary of Hart's theory:

According to Hart a rule that  $x$  ought to  $\phi$  when conditions  $C$  obtain exists in (the society)  $S$  if, and only if, the following conditions obtain:

- (1) Most  $x$ 's who are members of  $S$  regularly  $\phi$  when  $C$ . In other words, the rule is regularly complied with by members of the society to whom it applies.
- (2) On most occasions when an  $x$  does not  $\phi$  when  $C$ , he encounters some critical reaction from other members of  $S$ . In other words, deviations from the rule are the occasion for a critical reaction.
- (3) Such critical reactions do not themselves attract further criticism from members of  $S$ . Those who manifest critical reactions to deviations from the rule are not in turn subject to criticism for doing so, by members of  $S$ .
- (4) Members of  $S$  use expressions such as 'an  $x$  ought to  $\phi$  when  $C$ ' and 'it is a rule that an  $x$  ought to  $\phi$  when  $C$ ' to justify their own actions and to justify demands made of others or criticism of their behaviour.<sup>21</sup>

Raz is elsewhere clear that (2) to (4) are manifestations of the relevant critical reflective attitude, which is what the existence of the rule ultimately depends on.<sup>22</sup> As Raz notes, Hart's theory could easily be adapted for personal rules. Roughly, the adapted theory would say that

---

<sup>21</sup> Raz *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press, 1990) 52-53. I have capitalized some of Raz's symbols. For a similar summary see: Shapiro, 'How Rules Affect Practical Reasoning' in Bruno Verbeek (ed), *Reasons and Intentions* (Ashgate 2008) 134.

<sup>22</sup> See, e.g., Joseph Raz, *The Concept of a Legal System* (2<sup>nd</sup> edn, OUP 1980) 146-147.

you have a personal rule that you ought to  $\phi$  if and only if you generally  $\phi$  and make appropriate demands and self-criticisms.

Raz makes a number of points against Hart's theory. For example, he thinks that (2) is too restrictive because it only addresses occasions on which someone actually deviates from the rule. 'In fact', he says, 'hypothetical or theoretical expressions of opinion and manifestations of attitude by the members of the group are also relevant to the existence of rules.'<sup>23</sup> That includes 'views expressed in discussions of the rightness or wrongness of certain courses of conduct, to the education of the young, to expressions of opinion in literature, etc.'<sup>24</sup> He thinks (3) is too restrictive because it requires that 'offenders must admit, in most cases, the wrongness of their own behaviour'<sup>25</sup>, whereas it ought to be enough that 'they tend to participate in the critical attitude to other offenders in similar circumstances'<sup>26</sup>. To these criticisms, we might add that the attitude in question normally leads you to comply with the rule on appropriate occasions – something that Hart implies but never makes explicit.<sup>27</sup>

---

<sup>23</sup> Raz, *The Concept of a Legal System* (n 21) 147-148 n 4. Cf. Peter Hacker, 'Hart's Philosophy of Law' in Peter Hacker and Joseph Raz (eds), *Law, Liberty, and Morality: Essays in Honour of HLA Hart* (OUP 1977) 14: 'The circumstances which justify asserting that an agent accepts a rule ... [include] circumstances in which the rule does not apply at all, but e.g. its rationale is under discussion, hypothetical cases are examined, and the agent's verbal behaviour manifests his attitude'.

<sup>24</sup> Raz, *The Concept of a Legal System* (n 21) 147-148 n 4. Hart seems to have partly anticipated Raz. For example, Hart suggests that when there is a social rule the members of a group 'strive to teach or intend to maintain' the pattern of conduct and that rules are 'taught and efforts are made to maintain them': *Concept of Law* (n 1) 56, 86. For discussion see: MacCormick, *HLA Hart* (n 13) 45.

<sup>25</sup> Raz *Concept of a Legal System* (n 21) 148 n 1.

<sup>26</sup> Raz *Concept of a Legal System* (n 21) 148 n 1.

<sup>27</sup> See, e.g., Hart, *Concept of Law* (n 1) 58: 'The acceptance ... of ... a rule [requiring obedience to Rex or his successors] will be manifested during Rex I's lifetime in part by obedience to him'. See also: MacCormick, *HLA Hart* (n 13) 45: '[P]eople (when acting in the relevant circumstances) display or evince their intention to maintain the pattern by actually shaping what they do in accordance with the pattern, at least whenever they become conscious of the eligibility of some alternative'.

### 4.3 WARNOCK'S OBJECTION

The more significant objection to the practice theory (as Raz summarizes it) is that the theory is overinclusive. Conditions (1) to (4) are usually met when the members of a society generally do what they believe ought<sup>28</sup> to be done. However, often the members of a society generally do something, which they believe ought to be done, even though no rule in that society requires them to do so. So, Hart's theory often says that a society has a rule when actually it does not.

This criticism of Hart's theory originated with Geoffrey Warnock. Warnock gives this well-known example illustrating the problem:

[C]onsider the situation of the spectator of a cricket-match, ignorant of the game, and trying to work out what rules the players are following. He will find for instance that, when six balls have been bowled from one end, the players regularly move round and six balls are then bowled from the other end; deviations from this, he will observe, are adversely criticized. He will probably find also that, when a fast bowler is replaced by a slow one, some persons who were previously stationed quite close to the batsman are moved further away, some, probably, a lot further away; and he will find that, if this is not done, there is adverse criticism.<sup>29</sup>

The players tend both to move round after six balls have been bowled and to move near a slow batsman. In both cases the players react critically to deviations from the pattern. We might add that in both cases these critical reactions are normally expressed in normative terms (e.g., 'you ought to have moved further from the batsman') and are not themselves the subject of criticism. Yet there is a different diagnosis for each case.

[If the spectator] concludes that, in so acting, the players are following rules, he will of course be right in the first case and wrong in the second. There is no rule that a slow bowler should not operate with exactly the same field-setting as a fast one; this is indeed scarcely ever done, and it would nearly

---

<sup>28</sup> For present purposes it is not necessary to specify whether the 'ought' is all-things-considered or relative. The distinction figures prominently in the next chapter, however.

<sup>29</sup> Warnock, *Object of Morality* (n 20) 45-46.

always be regarded as wrong to do it, but that is because, quite independently of any rules, it is something which there is nearly always good reason to do.<sup>30</sup>

The conditions of Hart's theory are met in both cases, but there is a rule only in the first case.<sup>31</sup> The source of the problem is that a generally held belief that an action ought to be performed shows itself in convergent behaviour, critical reactions, and the use of normative language, just as the relevant critical reflective attitude does.<sup>32</sup> I will call this 'Warnock's Objection'.

Warnock's example is of a group situation, but the same objection holds in personal contexts. For example, I believe that Earl Grey tea tastes better without sugar. Usually I do not add any sugar. Sometimes I forget, though, and when I do I mildly criticize myself in the hope of remembering better next time. Yet I have no rule that prohibits me from adding sugar to Earl Grey tea. I would never describe adding sugar as a 'violation' or 'breach' of a rule, even to myself. Here is an example of Raz's that illustrates the same point:

Jack believes that he ought to read all of Iris Murdoch's novels and he does usually read them not long after publication. If he fails to read one of her novels within a year of its publication he tends to reproach himself for the omission. Yet he does not think of himself as having a rule that he should read all her novels. He does have other rules. He is a vegetarian and he cleans his teeth every evening. He does this because he believes that these are good rules to have. But he does not read Murdoch because of a belief in any rule.<sup>33</sup>

As in Warnock's cricket example, Jack believes that he ought to read all of Iris Murdoch's novels, but he has no corresponding rule.

---

<sup>30</sup> Warnock (n 20) 46. For variants on the example, and discussion, see: Shapiro, 'How Rules Affect Practical Reasoning' (n 21) 133-135; Raz, *Practical Reason and Norms* (n 21) 56.

<sup>31</sup> You might agree with Warnock that there is a rule that cricket players are to move to the opposite end after six balls have been bowled, but be less confident about his assertion that there is *not* a rule that they are to draw near a slow batsman. Certainly that is not one of the official rules of cricket. And, if it is a rule, it seems to have a different, less "strict" character than the rule about moving to the opposite end. It may, however, be what I later term a "rule of thumb". In that case, the significance of Warnock's example would be to show that Hart's theory does not mark the difference between strict rules and rules of thumb. The resolution of the issue is not crucial to the argument in the main text, because there are other examples that more clearly support Warnock's criticism. I mention some of them in the main text, and Raz provides others in *Practical Reason and Norms* (n 21) 55-56.

<sup>32</sup> There is one difference. When there is a rule, people sometimes say things like 'it is a rule that ...'; not so if they merely believe there is a conclusive reason.

<sup>33</sup> Raz, *Practical Reason and Norms* (n 21) 55-56.

Warnock's Objection is more robust than it might seem. Suppose that we made some of the changes to Hart's theory that Raz suggests. Suppose, for instance, that we allow that the critical reflective attitude is manifested, not only in the ways Hart mentions, but also in 'hypothetical or theoretical' expressions of support for the maintenance of the relevant pattern of conduct, including in the 'education of the young'. Warnock's Objection would still stand, because people voice the belief that an action ought to be performed in theoretical and educational contexts. So, making the change Raz suggests would not distinguish the critical reflective attitude from a belief that an action ought to be performed (which is not to say that was Raz's aim). Likewise, making it explicit that the critical reflective attitude displays itself in behaviour consistent with the pattern of conduct would not help, because belief that a pattern of conduct ought to be maintained would normally lead to the same behaviour.

Where are we left? Hart realized that when a society has a rule its members have a distinctive attitude. The trouble is, Hart describes the attitude too unspecifically. He describes the attitude in terms of how it leads someone to speak and act. The belief that an action ought to be performed leads someone to speak and act in the same ways. Yet people may believe that an action ought to be performed without them having a rule that it ought to be performed. So really we know two things. First, if a society has a rule that  $As$  ought to  $\phi$ , enough members of that society have a certain attitude, which leads them to act as they would act were they to believe that  $As$  ought to  $\phi$ . It leads them to act as though they have that belief or as if they have that belief.<sup>34</sup> (We could also say, if a society has a rule, enough of its members have an attitude that leads them to act as if they believe that the rule ought to be complied with, or that the rule-proposition<sup>35</sup> is true. I will use all of these expressions.)

---

<sup>34</sup> To be clear, by saying a person has an attitude which disposes him to act as if he believes there is a reason for  $As$  to  $\phi$ , I do *not* mean that the person does not believe there is a reason for  $As$  to  $\phi$ . It is an open question whether he has that belief.

<sup>35</sup> By "rule-proposition", I mean the proposition that is the content of a rule. The proposition that black ought to be worn at a funeral is a rule-proposition if there is a rule that black ought to be worn at a funeral.

Second, you can possess the attitude in question without possessing the belief that  $\mathcal{A}$ s ought to  $\phi$ . The attitude does not 'entail'<sup>36</sup> this belief.

#### 4.4 MACCORMICK'S REPLY

Warnock's Objection shows that it is insufficient to describe the attitude on which the existence of a rule depends purely in terms of its manifestations. We need to find some other, less observable feature of that attitude, and particularly one that is not shared by every belief in a reason to comply with a rule. Neil MacCormick describes the problem:

Hart has been insufficiently subtle in differentiating between relevant attitudes. ... Hart's elucidation of *rules* is not radically mistaken, but is only incomplete. What it rightly does is direct us to the question: what are the *attitudes* to patterns of social acting which, together with some regularity in action (or 'behaviour'), must exist or be held by human beings for it to be true that for some group of human beings a *rule* exists? To answer that crucial question we must start from, but cannot finish with, the materials which Hart has furnished.<sup>37</sup>

MacCormick has a suggestion for how we should distinguish the attitude Hart was after from other attitudes. He imagines that there is a group of vegetarians who say that no one ought to eat meat. He then asks what distinguishes the attitude possessed by the vegetarians from the attitude necessary for the existence of a rule.

What distinguishes our imagined vegetarians is that they make no assumption that the pattern of behaviour they favour, not eating meat, is a conventional one or one common to members of the groups in which they move. They hold it to be a preferable pattern to the common one, and their volitional commitment to it is a commitment on a point of principle and hence not in any way conditional upon common or shared observance in a group.<sup>38</sup>

---

<sup>36</sup> Strictly speaking, the proposition that you hold the attitude in question does not entail the proposition that you believe  $\mathcal{A}$ s ought to  $\phi$ , but it is more convenient to write as if the relationship were between the mental states.

<sup>37</sup> MacCormick, *HLA Hart* (n 13) 46-47.

<sup>38</sup> MacCormick, *HLA Hart* (n 13) 48.

As I understand him, MacCormick is saying that, when a society has a rule, enough of its members have an attitude that manifests itself in the ways Hart described *and* that is held conditional on general compliance with that rule among the members of that society. MacCormick's vegetarians do not have this attitude because they do not 'assume' that people in their society generally do not eat meat.

If this is indeed what MacCormick is thinking, then he has an easy way to deal with Warnock's example. The attitude that leads cricket players to move to the other end after six balls have been bowled depends on cricket players generally doing so. The attitude that leads the players to draw near a slow batsman does not depend on any general practice of drawing near a slow batsman. Thus, even though the attitudes manifest themselves in similar ways, they can be distinguished by the conditions under which they are held. This seems fairly commonsensical. That is, it seems right to say that the tendency to draw near a slow batsman is independent of what other players do, in a way in which the tendency to move to the other end after six balls have been bowled is not.

Whatever its virtues as a reconstruction of Hart's theory, MacCormick's proposal builds on what I have argued is a mistaken understanding of rules. If the conclusion of the last chapter is correct, then sometimes a rule exists in a society even though people in that society do not generally do as it says. Remember the grypsmen: they have an admission rule (the 'baptism rule') that none of them comply with. And remember the Ik: they have a rule of reciprocity that they do their best to evade, perhaps successfully. These societies have these rules, even though their members do not, or may not, generally comply with them. So, the attitude that is essential for a rule's existence cannot depend on people generally doing as that rule says. That is not to say the attitude *never* depends on general compliance with the rule; it is to say it need *not always* depend on it.

In the Postscript to *The Concept of Law*, Hart revises his theory. He says: ‘Rules are conventional social practices if the general conformity of a group to them is part of the reasons which its individual members have for acceptance ...’.<sup>39</sup> (The reference to ‘acceptance’ is best understood as a reference to the critical reflective attitude. More on this below.) He then says that ‘[m]y account of social rules is ... applicable only to rules which are conventions in [that] sense. ... This considerably narrows the scope of my practice theory’.<sup>40</sup> On one reading, Hart is doing what MacCormick suggests and making clear that the attitude he is interested in assumes compliance with a rule. In that case, my objection to MacCormick’s proposal applies equally to Hart’s revision. On another reading, Hart is restricting his theory to rules a society has partly in virtue of an attitude that *does* depend on conformity to a rule. There is nothing wrong with such a restriction, but it means that Hart’s project (at least when he writes in the Postscript) is no longer my project. I am trying to identify the necessary and sufficient conditions for a society to have a rule, not for it to have a certain type of rule.

#### 4.5 ACCEPTANCE

Hart says that the existence of a social rule depends on people having the right kind of attitude, which he describes as a ‘critical reflective attitude’. Frequently, Hart calls that attitude ‘acceptance’ or ‘acceptance of a rule’.<sup>41</sup> In the Postscript he summarizes the practice theory this way:

The account I have given ... treats the social rules of a group as constituted by a form of social practice comprising both patterns of conduct regularly

---

<sup>39</sup> Hart, *Concept of Law* (n 1) 255.

<sup>40</sup> Hart, *Concept of Law* (n 1) 256.

<sup>41</sup> See the page references at n 43 and 45.

followed by most members of the group and a distinctive normative attitude to such patterns *which I have called 'acceptance'*. This consists in the standing disposition of individuals to take such patterns of conduct both as guides to their own future conduct and as standards of criticism which may legitimate demands and various forms of pressure for conformity.<sup>42</sup>

In this summary the critical reflective attitude is cast as an attitude called 'acceptance'. When commentators describe Hart's theory, they usually refer to the attitude as 'acceptance of a rule'.<sup>43</sup>

Why does it matter what Hart calls the attitude? It matters because it helps us figure out what Hart thinks are the attitude's features. When Hart discusses the existence of social rules, he talks about a 'critical reflective attitude' and describes its manifestations. The focus on manifestations is what led to Warnock's Objection. Later, though, Hart talks about 'acceptance of a rule' in other contexts (not just in connection with the existence of social rules).<sup>44</sup> If Hart is talking about the same attitude the whole time – and it seems he is – then by looking at what he says about the acceptance of a rule we might learn something about the attitude in addition to how it is manifested. The hope is that we will learn something that gives us a way to respond to Warnock's Objection.

Let me be clear about my priorities. Hart's discussion of acceptance is scattered. There is no neat summary of the features of acceptance; nowhere does Hart say 'this is what acceptance is'. Hart's terminology is also extremely volatile. Sometimes, as in the block quote at the beginning of this section, he refers to acceptance *simpliciter*; much more often he refers to acceptance of a rule; occasionally he refers to acceptance of a rule as a standard or as a

---

<sup>42</sup> Hart, *Concept of Law* (n 1) 255. Emphasis added.

<sup>43</sup> See, e.g., Hart, *Concept of Law* (n 1) 55, 57, 58, 59, 60, 61, 89, 90, 101, 102.

<sup>44</sup> See, e.g., Hart's the discussion of acceptance of a rule in connection with points of view and normative statements and in connection with the existence of a legal system: *Concept of Law* (n 1) 89-90, 102-103, 115-117.

guide to conduct.<sup>45</sup> The differences, if any, among these terms are never explained. My aim is not exegesis, however. I am not concerned, as far as this thesis goes, with providing a natural or faithful reading of Hart. I am interested only in whether what Hart says about acceptance can give us some ideas of how to respond to Warnock's Objection.

Given my priorities, two of Hart's remarks are especially interesting. The first remark is about *why* a person may accept a rule. Hart suggests that the acceptance of a rule 'may be based on many different considerations: calculations of long-term interest; disinterested interest in others; an unreflecting inherited or traditional attitude; or the mere wish to do as others do'.<sup>46</sup> Some of these reasons are practical reasons. None of them are part of the case for the truth of the proposition that the rule ought to be complied with. That is, none are what Raz would term 'truth-related reasons'<sup>47</sup> for believing that the rule ought to be complied with. It may be in your interest to accept a rule that black ought to be worn at a funeral and, if Hart is right, it is possible to accept that rule *for* that reason, but it is not a truth-related reason to believe that black ought to be worn at a funeral.

The second remark is about *how* someone may come to accept a rule. Hart seems to think it is possible to *voluntarily* accept a rule.<sup>48</sup> That is, he seems to treat the attitude as one that is, or can be, voluntarily acquired. It is not clear exactly what Hart has in mind. He is sometimes read as claiming that acceptance of a rule is or may be the product of a

---

<sup>45</sup> Hart refers to 'acceptance' *simpliciter* in *Concept of Law* (n 1) 255; to 'acceptance of a rule' in numerous places, e.g., *Concept of Law* (n 1) 55-61; and to acceptance of a rule as a standard in *Concept of Law* (n 1) 116-117. Sometimes he refers to acceptance of a standard: *Concept of a Law* (n 1) 116; 'Scandinavian Realism' (n 4) 238.

<sup>46</sup> Hart, *Concept of Law* (n 1) 203. See also: Hart HLA Hart, *Essays on Bentham* (OUP 1982) 153ff; Hart, *Concept of Law* (n 1) 257: 'For some rules may be accepted simply out of deference to tradition or the wish to identify with others or in the belief that society knows best what is to the advantage of individuals. These attitudes may coexist with a more or less vivid realization that the rules are morally objectionable'. I should add that Hart's aim in these passages is to deny that acceptance of legal rules must be based on *moral* reasons, a claim that some scholars have doubted. See, e.g., Joseph Raz, *The Authority of Law* (OUP 1979) 153; MacCormick, *HLA Hart* (n 13) 162; Philip Soper, 'Law's Normative Claims' in Robert P George (ed), *The Autonomy of Law* (OUP 1996) 215-220.

<sup>47</sup> Joseph Raz, 'Reasons: Practical and Adaptive' in David Sobel and Stephen Wall (eds), *Reasons for Action* (Cambridge University Press 2009)

<sup>48</sup> Hart describes how, for the existence of coercive power, 'some at least must voluntarily co-operate in the system and accept its rules', and later on the same page refers to those who 'accept the system voluntarily'. Hart, *Concept of Law* (n 1) 203.

‘deliberation decision’<sup>49</sup>, but that too is ambiguous. One option is that you can decide to take some means and in that way come to accept a rule; acceptance of a rule would then be under your *indirect* control. Another option is that you can decide to accept a rule and, without doing anything else, acquire that rule; in that case acceptance of a rule would be under your *direct* control. The distinction is relevant because belief is beyond your direct control, but at least sometimes under your indirect control.

These remarks suggest a possible response to Warnock’s Objection. Hart could admit that acceptance of a rule and belief that the rule ought to be complied with manifest themselves in the same ways, but claim that the two attitudes differ in other respects. Hart might argue that acceptance can be held for non-truth-related reasons, but that belief cannot. Or he might argue that acceptance is under a person’s direct control, unlike belief, which is only under a person’s indirect control. The idea, in both cases, is to distinguish the attitudes by appealing to features other than their motivational force. For either response to be convincing, though, more would have to be said, especially about the nature of belief.

To be clear, I do not mean to attribute either response to Hart. His remarks are not detailed or unambiguous enough to justify doing so. My point is simply that Hart *does* say more about the attitude in virtue of which a rule exists in a society than how it is manifested, and that there is at least the potential in those remarks for a reply to Warnock’s Objection. In Chapter 9, I propose a theory of the existence of rules in a group. At the heart of that theory is an attitude, also called acceptance, which is distinguished from belief by, among other things, its responsiveness to non-truth-related reasons and its susceptibility to direct,

---

<sup>49</sup> See, e.g., Alf Ross, ‘Review of *The Concept of Law*’ (1962) 71 *Yale Law Journal* 1185, 11189. Ross criticizes Hart on this point, writing that ‘most people will feel themselves bound by the social norms of the group without ever being conscious of any choice or decision’. But it is not clear that Hart thinks that in *most* cases acceptance is deliberately chosen. He could think that it is *sometimes* chosen.

voluntary control. If that theory is defensible, as I believe it is, then it would be charitable to read Hart's remarks as consistent with that theory.<sup>50</sup>

#### 4.6 THE INTERNAL POINT OF VIEW

Hart uses the internal aspect and the acceptance of a rule to distinguish between two 'points of view', which is part of his more general discussion of obligation. Hart wants to show that the 'characteristic use [of a statement of obligation] is not to predict [a hostile reaction to deviation] but to say that a person's case falls under ... a rule'.<sup>51</sup> The difference between a prediction of a hostile reaction and saying that person's case falls under a rule is marked, says Hart, by the difference between the internal and external aspect.<sup>52</sup> He then says:

When a social group has certain rules of conduct, this fact affords an opportunity for many closely related yet different kinds of assertion; for it is possible to be concerned with the rules, either merely as an observer who does not himself accept them, or as a member of the group which accepts and uses them as guides to conduct. We may call these respectively the 'external' and the 'internal points of view'.<sup>53</sup>

The internal point of view is the point of view of a *member* of a group that has a rule, who *accepts* the rule, and who *uses* the rule as a guide to conduct. The external point of view is the point of view of an *observer* of the group who does *not accept* its rules.

Hart draws a further distinction between types of external points of view. Someone who takes an 'extreme' external point of view will simply note regularities of conduct and criticism in a group. For him, 'deviations by a member of the group from normal conduct will

---

<sup>50</sup> I return to Hart's notion of acceptance in Section 7.3 and Section 9.3.

<sup>51</sup> Hart, *Concept of Law* (n 1) 89.

<sup>52</sup> Hart, *Concept of Law* (n 1) 88-89: 'The following contrast [i.e., the contrast between a prediction of hostile reaction and a statement of obligation] again in terms of the "internal" and "external" aspect of rules may serve to mark what gives this distinction its great importance for the understanding not only of law but of the structure of any society'.

<sup>53</sup> Hart, *Concept of Law* (n 1) 89.

be a sign that hostile reaction is likely to follow, and nothing more'.<sup>54</sup> '[H]e will miss out a whole dimension of the social life of those whom he is watching'<sup>55</sup> because he will fail to see 'the way in which the group regards its own behaviour'<sup>56</sup>. His point of view

cannot reproduce ... the way in which the rules function as rules in the lives of those who normally are the majority of society. These are the officials, lawyers, or private persons who use them, in one situation after another, as guides to the conduct of social life, as the basis for claims, demands, admissions, criticism, or punishment, viz., in all the familiar transactions of life according to rules.<sup>57</sup>

In short, one who takes the extreme external point of view will fail to account for the internal aspect of rules. The more moderate (but still external) point of view is that of an observer who, 'without accepting the rules himself, assert[s] that the group accepts the rules, and thus ... from the outside refer[s] to the way in which *they* are concerned with them from the internal point of view'<sup>58</sup>.

The external point of view is the point of view of an observer, but members of a society who are concerned with its rules only 'because they judge that unpleasant consequences are likely to follow violation'<sup>59</sup> are said by Hart to come close to possessing the external point of view. Correlating deviations from a rule with hostile reactions, they say things like ' "I was obliged to do it", "I am likely to suffer for it if ...", "You will probably suffer for it if ...", "They will do that to you if ..." '.<sup>60</sup> They *use* the society's rules without *accepting* its rules. Because they do not accept the society's rules, they will *not* say things like ' "I

---

<sup>54</sup> Hart, *Concept of Law* (n 1) 90.

<sup>55</sup> Hart, *Concept of Law* (n 1) 90.

<sup>56</sup> Hart, *Concept of Law* (n 1) 90.

<sup>57</sup> Hart, *Concept of Law* (n 1) 90.

<sup>58</sup> Hart, *Concept of Law* (n 1) 89.

<sup>59</sup> Hart, *Concept of Law* (n 1) 90.

<sup>60</sup> Hart, *Concept of Law* (n 1) 90.

had an obligation” or “You have an obligation” for these are required only by those who see their own and other persons’ conduct from the internal point of view’.<sup>61</sup>

Hart builds on the distinction between the internal and external points of view to classify kinds of legal statements. After setting out his theory of rules of recognition of a legal system – rules that, among other things, establish criteria of legal validity in a community – Hart says that the ‘use of unstated rules of recognition ... in identifying particular rules of the system is characteristic of the internal point of view’.<sup>62</sup> Those who use rules of recognition this way ‘manifest their own acceptance of them as guiding rules and with this attitude there goes a characteristic vocabulary’<sup>63</sup>, the simplest instance of which is the statement ‘It is the law that ...’. Such a statement is an *internal statement* because it ‘manifests the internal point of view and is naturally used by one who, accepting the rule of recognition and without stating the fact that it is accepted, applies the rule in recognizing some particular rule of the system as valid’<sup>64</sup>. In contrast, a statement of the form ‘In England they recognize as law ... whatever the Queen in Parliament enacts ...’ is an *external statement* because it is ‘the natural language of an external observer of the system who, without himself accepting its rule of recognition, states the fact that others accept it’<sup>65</sup>.

There are objections that can be raised against Hart’s distinction between the internal and external points of view and the related distinction between internal and external statements.<sup>66</sup> It is not clear, for instance, precisely how Hart thinks the distinction between

---

<sup>61</sup> Hart, *Concept of Law* (n 1) 90.

<sup>62</sup> Hart, *Concept of Law* (n 1) 102.

<sup>63</sup> Hart, *Concept of Law* (n 1) 102.

<sup>64</sup> Hart, *Concept of Law* (n 1) 102-103.

<sup>65</sup> Hart, *Concept of Law* (n 1) 103

<sup>66</sup> See, e.g., Dennis Patterson, ‘Explicating the Internal Point of View’ (1999) *SMU Law Review* 67; Stephen Perry, ‘Holmes versus Hart: The Bad Man in Legal Theory’ and Scott J Shapiro, ‘The Bad Man and the Internal Point of View’, both in Steven J Burton, ed, *The Path of Law and its Influence: The Legacy of Oliver Wendell Holmes, Jr* (Cambridge University Press 2000); Stephen Perry, ‘Hart on Social Rules and the Foundations of Law: Liberating the Internal Point of View’ (2006) 75 *Fordham Law Review* 1171, 1180-1182; Scott Shapiro, ‘What is the Internal Point of View?’ (2006) 75 *Fordham Law Review* 1157.

these points of view maps onto the distinction between the perspectives of a member versus a non-member of a group, or the distinction between the perspective of an observer and a non-observer. Another problem, which Raz raised<sup>67</sup> and Hart conceded<sup>68</sup>, is that statements like ‘You have an obligation’ are sometimes made by those who do not possess the internal attitude towards the relevant rule, and that statements like ‘It is the law that ...’ are sometimes made by those who do not possess the internal point of view towards the relevant rule of recognition. The details of these difficulties are not important for my purposes.

Of potentially greater relevance is Kevin Toh’s argument that Hart provides an ‘expressivist’ analysis of internal legal statements.<sup>69</sup> ‘Expressivism’ refers to theories that try to explain the meanings of normative terms by telling us what mental state a person expresses by uttering a statement containing those terms. The mental state expressed could be cognitive (e.g., a belief) or noncognitive (e.g., a desire), but usually expressivists are also noncognitivists. Alan Gibbard, for example, argues that in uttering statements of rationality, a person expresses his or her acceptance of a system of norms, and thus expresses a noncognitive mental state.<sup>70</sup> Theories like Gibbard’s are called ‘norm-expressivist’ theories.

Toh offers several reasons for thinking that Hart is an expressivist, including that he is motivated by some of the same concerns as expressivists and that the philosophical climate in the 1950s and 1960s was favourable to expressivism. Toh claims, more specifically, that Hart is, or is what we would now call, a norm-expressivist with regard to internal legal statements. Toh makes that attribution on the basis of several of Hart’s remarks, such as the

---

<sup>67</sup> See, e.g., Raz, *Practical Reason and Norms* (n 21) 172-173; *The Authority of Law* (n 46) 155.

<sup>68</sup> HLA Hart, *Essays in Jurisprudence and Philosophy* (OUP 1983) 14.

<sup>69</sup> Kevin Toh, ‘Hart’s Expressivism and his Benthamite Project’ (2005) 11 *Legal Theory* 75. See also: Kevin Toh, ‘An Argument Against the Social Fact Thesis (And some Additional Preliminary Steps Towards a New Conception of Legal Positivism)’ (2008) 27 *Law and Philosophy* 445; Kevin Toh, ‘Raz on Detachment, Acceptance and Describability’ (2007) 27 *Oxford Journal of Legal Studies* 403.

<sup>70</sup> This is Toh’s description of Gibbard’s theory, set out in *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (OUP 1990).

claim that those who utter internal legal statements ‘manifest their own acceptance of [the rules of recognition] as guiding rules’. On Toh’s view, Hart should be thought of as trying to explain the meaning of an internal legal statement by telling us that such statements expresses the speaker’s acceptance of the relevant rules of recognition.<sup>71</sup>

Can this help with my project? Certainly Toh is interested in acceptance and rules and so am I. And Toh’s careful discussion adds to the evidence, if any is needed, that Hart does not think that the ‘critical reflective attitude’ necessary for the existence of a social rule is a belief that the rule ought to be complied with. But the differences between our projects are more significant. Toh is interested in acceptance and rules only insofar as it helps him to understand what Hart really thought about certain legal statements. My concern is not how best to read Hart or how to understand legal statements; it is to find the best account of what it is to have a rule. Nor does Toh seem especially interested in the nature of acceptance. That is understandable, because his argument requires only that acceptance of a rule is a noncognitive attitude. But it is not enough for my purposes. Before I can show what attitude rule-possession consists in, I need to describe the attitude’s features and explain how it differs from a belief that a rule ought to be complied with – tasks with which Toh’s analysis does not seem to assist.

There is another reason why Toh’s analysis is unlikely to advance my project. I have spent much of this chapter explaining why, to understand what it is to have a rule, we need to go beyond what Hart says about acceptance. So, it is not helpful, for my project, to find out that someone thinks of acceptance as Hart does, because that would not add to what we know about that attitude. For Toh, things work the other way around. He wants to show that Hart is a norm-expressivist like Gibbard, which means showing that Hart and Gibbard think

---

<sup>71</sup> More fully, Toh says that Hart also understands the speaker to *presuppose* that the relevant rule of recognition is generally accepted and complied with in his community: Toh, ‘Hart’s Expressivism’ (n 69) 88.

of acceptance similarly. The less there is to distinguish their views the better, from Toh's perspective. And, indeed, Hart's and Gibbard's views *are* similar, especially in how they think acceptance of a rule or norm is displayed or manifested. Toh says:

Gibbard characterizes an acceptance of a norm involved in a statement of rationality as a set of dispositions to be governed by a norm, and to avow it in unconstrained normative discussion, as a result of the workings of mutual demands for consistency in the positions one takes in normative discussion. Hart has strikingly similar things to say about what an acceptance of a norm amounts to. Given this similarity, it can be thought that the difference [between the statements Gibbard and Hart are interested in] has to do *not* with the difference in the acceptances involved but instead with the difference in the norms accepted.<sup>72</sup>

Having shown that Hart and Gibbard use a very similar notion of acceptance, Toh is better positioned to conclude that Hart is a norm-expressivist. By the same reasoning, I have little to gain from Gibbard's discussion. The better Toh's argument goes for him, the less relevance it has for me.

#### 4.7 CONCLUSION

Hart set two conditions for a society to have a rule. The members of that society must generally comply with that rule, and enough of them must possess a certain attitude. I earlier rejected the first condition, leaving only the second condition. That a society has a rule simply means that enough of its members have the right attitude. Thus, the answer to the thesis question, 'what is it to have a rule?', essentially becomes, 'what attitude is the right attitude?' Thanks to Hart and his critics, we know that this attitude leads you to act as if you believe that the rule ought to be complied with, and that it is possible to have this attitude without having that belief. In the next two chapters, I try to find out more about attitude in question.

---

<sup>72</sup> Toh, 'Hart's Expressivism' (n 69) 89. Citations omitted.

# 5 Strength and Strictness

---

## 5.1 INTRODUCTION

We know that for a society to have a rule is for enough of its members to have a certain attitude, and the task now is to figure out what attitude it is. In this chapter I focus on the effect of having a rule on deliberation (Section 2), and the relation between having a rule and certain beliefs (Section 3). I draw on what I find to distinguish four types of rules, along two dimensions: strength and strictness. I conclude this chapter knowing substantially more about the features of the attitude in which rule-possession consists. That sets the agenda for later chapters by focusing the inquiry on attitudes with those features.

## 5.2 STRENGTH

### *The Peremptory Effect*

Consider one way to make a decision. You try to identify the reasons for and against the available actions. You balance those reasons against each other to determine which reasons defeat which others. Eventually you arrive at a judgment of what you ought to do all things considered. That judgment leads you to decide to perform that action. Now, this is a highly simplified picture of just one form of decision-making. All that matters, for my purposes, is that something like it describes an ordinary process of making a decision. Call the process it

describes ‘ordinary decision-making’ and the first step (when it is part of the whole) the ‘balancing process’.

When people have a rule, and they judge that it applies, they typically do not decide whether to comply with it in the ordinary way. In particular, they tend not to go through the full balancing process. Instead, they proceed as though what ought to be done was resolved, at least in certain respects, making further deliberation idle. Having a rule thus has a *peremptory effect* on deliberation.<sup>1</sup> When setting the table, for instance, you do not weigh all the pros and cons of putting the fork on the left or the right. You proceed as though it is settled that the knife ought to go on the right.<sup>2</sup> And it is not just deliberation about whether to comply with a rule that is cut short; it is also deliberation about whether to demand compliance with that rule, to criticize its violation, and so on. As Raz says, ‘it must be admitted that for the most part the presence of a norm [or rule] is decisive’.<sup>3</sup> ‘Once it is established that a norm applies to the case at hand we need not be concerned with the weights of the conflicting reasons’.<sup>4</sup>

Having a rule leads people to skip or shorten the balancing process, but that behaviour is not unique to those who have a rule. You go through the balancing process to arrive at a judgment of what you ought to do. Normally, you skip that process entirely if you *already* believe that you ought to perform some action, all things considered, because there is little point in inquiring about what you already know. Likewise, normally, you do not bother

---

<sup>1</sup> [‘T]he word “peremptory” in fact just means cutting off deliberation, debate, or argument and the word with this meaning came into the English language from Roman law, where it was used to denote certain procedural steps which if taken precluded or ousted further argument.’ HLA Hart, ‘Commands and Authoritative Legal Reasons’ in *Essays on Bentham* (OUP 1982) 253-254. I use the word ‘peremptory’ in the way Hart defines it, but I do not mean to imply any more significant connection between what I call the peremptory effect of having a rule and Hart’s discussion of peremptoriness.

<sup>2</sup> There are many other examples. See, e.g., Alan Goldman, *Practical Rules: When We Need Them and When We Don’t* (Cambridge University Press, 2002) 28: ‘Having a mandatory retirement rule ... removes the responsibility and the need for employers and employees to engage ... in weighing all the factors that might be relevant positively or negatively to the decision on whether to retire’.

<sup>3</sup> Joseph Raz, *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990) 79.

<sup>4</sup> Raz, *Practical Reason and Norms* (n 3) 79. See also: Scott Shapiro, ‘On Hart’s Way Out’ (1998) 4 *Legal Theory* 469, 493: ‘When someone who is guided by a legal rule is confronted with a situation where the rule applies the person does not deliberate about whether to follow the rule – he or she simply follows the rule’.

weighing reasons of a certain type if you already believe that you ought to perform some action, relative to those reasons. You turn your attention to your other reasons instead.

I can now add some detail to the conclusion in the last chapter. I said that a society has a rule only if enough of its members act as if they believe it ought to be complied with. Now we can see that the character of the ‘ought’ may vary. In some cases, it may be an all-things-considered ‘ought’. In other cases, the ‘ought’ may be relative to a type of reason.

### *Strong rules and weak rules*

When having a rule leads the members of a society to act as if they believe it ought to be complied with, all things considered, I will say the rule is a ‘strong rule’. It seems possible, for example, that constitutional actors act as if they believe that the Queen ought to assent to legislation no matter what. In that case, the convention that she ought to do so would be a strong rule.<sup>5</sup> Having a rule may, instead, lead people to act as if they believe it ought to be complied with, relative to some type of reason (e.g., reasons of convenience), while leaving it open what ought to be done in an all-things-considered sense. These rules are ‘weak rules’. Having a strong rule means skipping the balancing process; having a weak rule means shortening it.

At one time the claim that a rule can be weak might have been controversial. In his ‘Model of Rules I’, Ronald Dworkin seems to suggest that rules are always strong. He writes: ‘Rules are applicable in an all-or-nothing fashion. If the facts stipulated by the rule are given, then either the rule is valid, in which case the answer it supplies must be accepted, or it is not,

---

<sup>5</sup> In Britain the monarch has not refused assent in over 300 years. Sometimes people argue that the Queen ought to refuse assent if legislation is plainly and deeply evil – e.g., it authorizes the torture of babies – or egregiously unconstitutional – e.g., it suspends elections indefinitely and without justification. I am not rejecting these examples. I am only assuming that this *could* be a strong rule. For other examples, see Section 9.2.

in which case it contributes nothing to the decision'.<sup>6</sup> Dworkin allows that a rule may have an exception, but he thinks that 'an accurate statement of the rule would take this exception into account, and any that did not would be incomplete'.<sup>7</sup> Rules are dispositive or inapplicable, and it is always possible to say in advance which it will be.

Dworkin contrasts his notion of a rule with what he called principles. Principles have a 'dimension of weight' that rules lack, according to Dworkin. Principles may be relevant but not dispositive. They count in favour of an action, but not necessarily conclusively.

A principle like 'No man may profit from his own wrong' ... states a reason that argues in one direction, but does not necessitate a particular decision. ... [O]ur principle may not prevail, but that does not mean it is not a principle of our legal system, because in the next case, when these contravening considerations are absent or less weighty, the principle may be decisive. All that is meant, when we say that a particular principle is a principle of our law, is that the principle is one which officials must take into account, if it is relevant, as a consideration inclining in one direction or another.<sup>8</sup>

Dworkin adds that a full statement of a principle cannot capture all 'counter-instances' – the cases in which a principle is relevant but not conclusive. 'They are not, even in theory, subject to enumeration', he claims, because they would have to include 'those numberless imaginary cases in which we know in advance that the principle would not hold'.<sup>9</sup> Thus, principles, unlike rules, may be relevant but inconclusive, and it is impossible to state in advance when that will be.

Dworkin's distinction between rules and principles generated a great deal of discussion in legal philosophy.<sup>10</sup> My purpose is not to try to reopen that debate. Suffice it to

---

<sup>6</sup> Ronald Dworkin, *Taking Rights Seriously* (Harvard University Press, 1978) 24; see also 71-80.

<sup>7</sup> Dworkin, *Taking Rights Seriously* (n 6) 25.

<sup>8</sup> Dworkin, *Taking Rights Seriously* (n 6) 26.

<sup>9</sup> Dworkin, *Taking Rights Seriously* (n 6) 25.

<sup>10</sup> For a small sample, see: Joseph Raz, 'Legal Principles and the Limits of Law' (1972) 81 *Yale Law Journal* 823; Colin Tapper, 'A Note on Principles' (1971) 34 *Modern Law Review* 628, 628-632; Hart (n 8) 259-268; Larry Alexander and Ken Kress, 'Against Legal Principles' (1997) 82 *Iowa Law Review* 739; Frederick Schauer, 'Prescriptions in Three Dimensions' (1997) 82 *Iowa Law Review* 911; Timothy Endicott, 'Are There Any Rules?' (2001) 5 *Journal of Ethics* 199.

say that few philosophers of law now endorse the distinction between rules and principles as Dworkin draws it. In part, that is because the distinction does not correspond very well to our ordinary use of the terms. As Raz says, ‘we do not normally use the rule/principle distinction to mark the difference between *prima facie* and conclusive reasons or between the standards which establish them’.<sup>11</sup> Schauer gives some examples that demonstrate the point:

When I drive in excess of a precise speed limit in order to rush an injured child to the hospital, or when the observant Jew eats pork in order to avoid starvation, the force of the applicable rule has been overridden by more exigent considerations. Surely these are rules, if anything is, and just as surely rules are routinely overridden in circumstances comparable to those just mentioned. In these and countless other instances, the reason for action supplied by an applicable rule is not in the particular circumstances sufficient to resist the reasons for action supplied by other considerations, some of which may but need not be other rules.<sup>12</sup>

Schauer’s examples are best seen as examples of weak rules. Observant Jews, for instance, act as if they believe that, relative to reasons of pleasure and the like, but not relative to reasons of survival, pork ought not to be eaten. That is to say, they act as if they believe there is a reason against eating pork that is not defeated by reasons of pleasure, but which may be defeated by reasons of survival. It is not possible to list in advance all survival-related reasons, making this rule – and weak rules generally – open-ended in the way Dworkin described principles as being. (Whether weak rules are principles, or whether there is a residual distinction between rules and principles, is not an issue I will discuss.)

Rules of etiquette also provide examples of weak rules. Consider the rules for the use of knives and forks from a 19<sup>th</sup> century guide to etiquette, *Modern Etiquette in Private and Public*.

---

<sup>11</sup> Raz ‘Legal Principles and the Limits of Law’ (n 10) 82.

<sup>12</sup> Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (OUP 1991) 115. See also: Frederick Schauer, ‘On the Supposed Defeasibility of Legal Rules’ (1998) 51 *Current Legal Problems* 223, 228-230; William Twining and David Miers, *How to do Things with Rules* (5<sup>th</sup> edn, Cambridge University Press 2010) 134-135. For a particularly good example, see: Jon Elster, *The Cement of Society: A Study of Social Order* (Cambridge University Press 1989) 109: ‘When the survivors of a 1972 air crash in the Andes ate the flesh of the dead, they were not condemned but forgiven, absolved and even turned into celebrities of heroic stature’.

Fish follows the soup. You must eat it with a fork, unless silver knives are provided. Break a little crust off your bread, to assist you in taking up your fish, but it is better to eat with the fork only, which you may do if it be turbot or salmon.

...

After soup and fish come the side dishes, as they would be called, if they were on the table – the oyster or lobster patties, quenelles, etc.

Remember, that for these you use *the fork only*; as, indeed, you should for all dishes which do not absolutely require a knife. You must use a knife, of course, for cutlets of any kind, although they *are* side-dishes. It is proper to eat all soft dishes, as mince, etc., with the fork only.

Do not put your hands on the table, except to eat or carve (the latter is not required at a dinner *à la Russe*). Do not use your handkerchief if you can help it; if you *must* do so, let it be as inaudibly as possible.

...

For the partridge or pheasant, of course you use the knife and fork; all sweets are eaten with the fork, or spoon and fork, as you like; but the spoon is only required for cherry-tart, or anything of that nature, custard, etc.

Ladies scarcely ever eat cheese after dinner. If eaten, remember it should be with *a fork*, not a knife.

You should *never*, by any chance, put a knife near your mouth.<sup>13</sup>

These rules, like most rules of etiquette, are weak rules, not strong rules. People act as if fish ought to be eaten with a fork even if it is more convenient or practical to do otherwise. But they do not act as if fish ought to be eaten with a fork *no matter what*, as they would were the rule a strong rule. Few people would voice disapproval if a seriously ill person broke one of these rules by, for example, putting a knife near his or her mouth. And, even within the bounds of convenience or practicality, rules of etiquette sometimes yield. The anonymous writer of *Modern Etiquette* acknowledges as much when he or she says that ‘you should [use a fork] for all [side] dishes which do not absolutely require a knife’. If the convenience-related

---

<sup>13</sup> Anonymous, *Modern Etiquette in Private and Public* (Frederick Warne 1872) 16-18 excerpted in Mary Douglas (ed), *Rules and Meanings: The Anthropology of Everyday Knowledge* (Routledge 1973) 217-218.

reasons for using a knife are compelling enough – if you *absolutely* need a knife to eat a dish – the rule is overcome.

In saying that rules of etiquette and manners are usually weak rules, I do not mean they are *easily* overridden, or that some people do not treat them as strong. Emily Post, author of several famous books on etiquette and manners, views at least some rules of manners as being strong, or nearly so. She gives this example:

[P]arents must never disagree before the children. It simply can't be! Nor can there be an appeal to one parent against the other by a child.

‘Father told me to jump down the well!’

‘Then you must do it, dear,’ is the mother’s only possible comment. When the child has ‘jumped down the well,’ she may pull him out promptly, and she may in private tell her husband what she thinks about his issuing such orders and stand her own ground against them; but so long as parents are living under the same roof, that roof must shelter unity of opinion, so far as any witnesses are concerned.<sup>14</sup>

Perhaps Post was exaggerating for effect, or perhaps she was alone in her view of the importance of the rule regarding parental disagreements in front of children. Otherwise, it serves to demonstrate that not all rules a society has are ones it should have.

Weak rules of etiquette yield to reasons of safety and health and the like. Some rules do not yield to such reasons, but they are weak nonetheless. The practice of dueling in the American South provides an example.<sup>15</sup> The origin of a duel was an insult, or perceived insult, of one gentleman’s honour by another.<sup>16</sup> The insult could be trivial; satisfaction could still be demanded. Once a challenge was issued, it had to be accepted, with very few exceptions

---

<sup>14</sup> Emily Post, *Etiquette in Society, in Business, in Politics and at Home* (Funk and Wagnalls Company 1922) 592.

<sup>15</sup> On duelling in the American South generally, see: Warren Schwartz, Keith Baxter, and David Ryan, ‘The Duel: Can These Gentlemen Be Acting Efficiently?’ (1984) 13 *Journal of Legal Studies* 321; Lawrence Lessig, ‘The Regulation of Social Meaning’ (1995) 62 *University of Chicago Law Review* 943, 968-972; Jack Williams, *Dueling in the Old South: Vignettes of Social History* (Texas A&M University Press 2000); Harwell Wells, ‘The End of the Affair: Anti-Dueling Laws and Social Norms in Antebellum America’ (2001) 54 *Vanderbilt Law Review* 1805.

<sup>16</sup> It was crucial that both the person making and the person receiving the challenge were gentlemen. See Williams, *Dueling in the Old South* (n 15) 27: ‘No gentleman ever accepted a challenge from one not considered his social equal.’

(more about which in a moment). Refusal resulted in severe social disapproval and a loss of honour, despite the substantial risks involved.

The combatants would stand at a certain distance apart and fire pistols at one another. If the insult was minor, the duel might end with a single shot and no injuries. But if the insult was serious, the duel would continue until one of the duelists was injured or killed. The chance of injury was estimated at one in six, and the chance of death at one in fourteen.<sup>17</sup> The risks led all Southern states to prohibit duelling, with heavy penalties imposed on combatants and their seconds.<sup>18</sup>

Even though it was dangerous and illegal, dueling persisted for centuries in the American South. Gentlemen risked death or imprisonment to avenge seemingly insignificant social slights, which is consistent with a belief that a duel ought to be accepted *whatever* the costs. It might appear, then, that the rule requiring the acceptance of a challenge was a strong rule. Interestingly, that was not the case. A gentleman could legitimately decline a challenge to a duel if there were compelling *honour-related* reasons to do so, as there sometimes were. For example, although one legal sanction for dueling was a term in prison, another was a ban on holding public office.<sup>19</sup> Holding public office, at that time, for this group, was not merely desirable. As Lawrence Lessig writes, '[h]olding public office ... was itself a duty of the elite. Thus, exclusion [from public office] created a conflict in the duties faced by the elite, and hence an elite-based reason for refusing the challenge of a duel'.<sup>20</sup>

---

<sup>17</sup> A Traveller, *The Art of Dueling* (Willoughby 1836) in Charles Haven, *Shooting Muzzle Loading Handguns* (Guns Inc 1947).

<sup>18</sup> The laws were quite carefully designed. They did not require anyone to be harmed; they set substantial penalties, but not so substantial as to lead to jury nullification; and they disqualified *any* person from participating in a duel from holding public office. See: Schwartz and others, 'The Duel' (n 15) 326.

<sup>19</sup> Schwartz and others, 'The Duel' (n 15) 326; Lessig, 'The Regulation of Social Meaning' (n 15) 971-972; Wells, 'The End of the Affair' (n 15) 1825-1837.

<sup>20</sup> Lessig, 'The Regulation of Social Meaning' (n 15) 971. Lessig brings out the point in this imagined exchange between two potential combatants (972):

Challenged: 'I decline your challenge on the grounds that it will bar me from serving in public office in the future.'

The rule requiring gentlemen to accept a challenge to a duel was a weak rule, on my classification scheme. Unlike most weak rules, however, that rule did not yield when a person's health or life is at stake. It trumped what would normally be thought of as the weightiest reasons to decline a duel, while yielding to considerations of honour. It was the opposite of rules of etiquette and manners, which yield to serious conflicting considerations, but overcome less serious ones.

### *Efficiency and Peremptoriness*

Their peremptory effect is part of what makes having a rule useful.<sup>21</sup> Completing the balancing process can be costly. It can take time and energy to identify and weigh reasons, or to learn how to do so. It can be tiring, unpleasant, and stressful. By helping to circumvent the balancing process, having a rule can help people to avoid or reduce these costs. It helps improve efficiency by providing a shortcut to action. Schauer puts the point well:

[T]he observant Jew is relieved by the rules of Kashruth from having to train herself as a biologist just to know which foods to eat and which to avoid. And in cultures in which a particular dress is mandatory (e.g. grey flannel suits for business; gowns at high table; no white shoes before Memorial Day), a common argument for those rules is precisely that they eliminate the calculations, the anguish, and the expenditures that would otherwise be necessary, thereby freeing time, money, and mental space for more worthwhile endeavors.<sup>22</sup>

---

Challenger: 'But you are obligated as a gentleman to accept my challenge.'

Challenged: 'I am also obligated as a gentleman to serve my state, and I consider that duty superior to my duty to give you satisfaction. I agree it would be better if I were free to satisfy both duties, and I would willingly accept your challenge if it did not disable me from serving my state. I have no control over that, however. I must, therefore, respectfully decline.'

<sup>21</sup> John Stuart Mill, *System of Logic, Ratiocinative and Inductive* (first published 1843, Kessinger Press 2004) 617-618: '[Rules] point out the manner in which it will be least perilous to act, where time or means do not exist for analysing the actual circumstances of the case, or where we cannot trust our judgment in estimating them'. See also, Raz, *Practical Reason and Norms* (n 4) 59-62; Schauer, *Playing by the Rules* (n 12) 145-155.

<sup>22</sup> Schauer, *Playing by the Rules* (n 12) 146-147.

In all these cases, having a rule is a simpler alternative to ordinary decision-making.<sup>23</sup> Rather than spend your time and energy weighing reasons, you can simply look to the rule to decide what to do.<sup>24</sup>

Raz gives an example of a personal rule that reduces deliberation costs. He imagines that Jill wishes to save herself the hassle of weighing the pros and cons of various holiday destinations year after year, and relies on a rule to help her do so.

[Jill] faces many conflicting considerations, but she has no intention to act on the balance of reasons. She has adopted a rule to spend her holidays in France and she did so precisely in order to spare herself the necessity of deciding every year what to do during the holidays. ... When the holidays come she is not going to reconsider the matter. Her mind is already made up.<sup>25</sup>

That Jill has this rule means that she will act as if she believes she ought to go to France on her holidays. She might act as if she believes that she ought to go to France, all things considered, or just relative to reasons of convenience, pleasure, cost, etc. Either way, she will not go through the full balancing process when it comes time to decide where to go this year or next year, because 'her mind is already made up'.

That people have rules in order to simplify their decision-making is significant for two reasons. First, it is added evidence that having a rule has a preemptory effect on deliberation. People would not rely on rules to simplify their decision-making unless rules were capable of doing so. Second, it suggests that the attitude in which having a rule consists is shaped (at

---

<sup>23</sup> Determining whether a rule applies is not necessarily a straightforward matter. Indeed, making that determination correctly could potentially be more costly and susceptible to error than the weighing process. In that case, having the rule may not be instrumentally justified.

<sup>24</sup> Goldman, *Practical Rules* (n 2) 34: '[T]he application of rules saves time, effort, and investigative and cognitive resources because of their simplifying function'; Raz, *Practical Reason and Norms* (n 3) 59: 'When a situation to which it applies ... the norm subjects can rely on the rule, thus saving much time and labour'; Cass Sunstein, 'Political Conflict and Legal Agreement' (The Tanner Lectures on Human Values, delivered at Harvard University 1994) <<http://www.tannerlectures.utah.edu/lectures/documents/Sunstein96.pdf>> accessed 7 July 2011, 214-215: 'Without rules, decisions are extremely expensive; rules can produce enormous efficiency gains. Every day we operate as we do because of rules, legal and nonlegal, and often the rules are so internalized that they become second-nature, greatly easing the costs of decisions and making it possible to devote our attention to other matters.'

<sup>25</sup> Joseph Raz, 'Reasons for Action, Decision, and Norms' (1975) 84 *Mind* 481, 497-498. Raz discusses the example again in 'Facing Up: A Reply' (1989) 62 *Southern California Law Review* 1153, 1156-1157.

least partly) by practical considerations, and that people acquire the attitude *for* practical reasons. In the last chapter I said that Hart claimed that this attitude is shaped by practical reasons, so now I have evidence in support of Hart's claim. I return to this point in chapters 9 and 10.

### 5.3 STRICTNESS

A society has a rule only if enough of its members act as if they believe it ought to be complied with. So far, I have been exploring the relationship between the character of the 'ought' and the type of rule possessed. Let me now turn to the relationship between a society having a rule and its members' actual beliefs. I said in the last chapter that the members of a society may believe that *As* ought to  $\phi$  without having a rule that *As* ought to  $\phi$ . (Nothing changes if we specify that the 'ought' is all-things-considered or relative to a type of reason. Warnock's cricket example could be adapted to prove the point, as could Raz's example of Jack, the Iris Murdoch fan.) What about the converse case? Is it possible for a society to have a rule that *As* ought to  $\phi$  without its members believing that *As* ought to  $\phi$ ? The question is not as straightforward as it may seem, largely because of the complexities surrounding the ways reasons may defeat each other. I address the question in the next chapter, where I give it a qualified answer. For now, I will lay the groundwork by considering the relationship between having a rule and beliefs about the balance of reasons.

## *Strict Rules and Rules of Thumb*

You can have a rule without believing that the balance of reasons always favours complying with that rule.<sup>26</sup> Hart makes this point when he describes how, on taking office, a judge finds a practice ‘according to which any judge of the system is required to apply in the decision of cases the laws identified by specific criteria or sources’.<sup>27</sup> Judges are disposed to comply with that rule, Hart says, whatever their views of the merits of doing so.

The judges not only follow this practice as each case arises but are committed in advance in the sense that they have a settled disposition to do so without considering the merits of so doing in each case and indeed would regard it not open to them to act on their view of the merits. ... [A judge’s] view of the merits may be favourable or unfavourable, or simply absent ....<sup>28</sup>

Judges and other officials have a rule that requires them to apply the law, but they may or may not believe that the balance of reasons (the ‘merits’) always favours applying the law. They are disposed to apply the law either way.

Not only can you have a rule without believing the balance of reasons favours compliance with it; you can have that rule and believe that, in *this* case, the balance of reasons does *not* favour compliance. What happens when you do marks the difference between two important types of rules. It may be that, despite believing that the balance of reasons does not favour compliance with the rule in the circumstances, having the rule *still* leads you to act as if

---

<sup>26</sup> See, e.g., Geoffrey Warnock, *The Object of Morality* (Methuen & Co Ltd 1971) 36: ‘A person is acting in compliance with a rule ... if he supposes that he (or one) is to act in that way, whether or not in every case there is, there and then, good reason to do so’; Raz, *Practical Reason and Norms* (n 3) 75: ‘[E]veryone who follows a rule’ may think he should perform the action but have the ‘characteristic mixed reaction to the effect that doing so is not all right, i.e., that there is a different calculation according to which he ought not to perform the action’; Joseph Raz, ‘Promises and Obligations’ in Peter Hacker and Joseph Raz (eds), *Law, Morality and Society: Essays in Honour of HLA Hart* (OUP 1977) 223-224; Ronald Heiner, *Imperfect Decisions and the Law: On the Evolution of Legal Precedent and Rules* (1986) 15 *Journal of Legal Studies* 227, 236; Raz, *Practical Reason and Norms* (n 4) 75; Schauer, *Playing by the Rules* (n 4) 4; Robert Nozick, *The Nature of Rationality* (Princeton University Press 1993) 10, 27; Scott Shapiro and Edward McClennen, ‘Rule-Guided Behaviour’ in Peter Newman (ed), *The New Palgrave Dictionary of Economics and the Law* (Palgrave Macmillan 1998) 366.

<sup>27</sup> HLA Hart, *Essays on Bentham* (OUP 1982) 158.

<sup>28</sup> Hart, *Essays on Bentham* (n 27) 158-159. Hart is responding to the claim that the judges of a legal system must believe or pretend to believe in the moral justifiability of the law.

you believe it ought to be complied with. Such rules have a binding or, as I will say, ‘strict’ character. Warnock gives an example of a strict rule of cricket:

Consider ... an actual rule of cricket – say that six balls, and no more, are to be delivered from each end in turn. ... [U]mpires do not, on each occasion when bowlers have bowled six balls, bring the over to an end because they see, on each occasion, good reason to do so. There are in fact plenty of cases in which, from the batsman’s or bowler’s or even spectator’s point of view, it would be an excellent thing to have more than six balls; but the fact is that the merits of such cases do not come into it.<sup>29</sup>

Umpires do not bring the over to an end because they believe the balance of reasons – the ‘merits’ of the case – favour doing so. Sometimes, the umpire is confident that the balance of reasons favours continuing on. He or she brings the over to an end anyway. The ‘question is removed from the sphere of judgment on the particular merits of each case’.<sup>30</sup> In Hart’s presentation, the rule of recognition is also strict, because judges have a ‘settled disposition’ to apply the law even when they believe the merits oppose doing so.

There is a second possibility. It may be that having a rule leads you to act as if you believe it ought to be complied *except* when you believe, to a certain degree of confidence, that in this case the balance of reasons favours doing something else. Such rules are not strict, but they vary depending on the degree of confidence it takes for the rule to yield. At one extreme is a rule that yields to a belief of *any* degree of confidence that the balance of reasons does not favour compliance. At the other extreme is a rule that yields only to a flat-out belief. In between would be rules capable of being displaced by beliefs of intermediate degrees of confidence. When the requisite degree of confidence is met, whatever it is, you set aside the rule and go through the normal weighing process.

Most of the discussion in the literature is about strict rules, but rules that are not strict are probably more common. They include recipes, style guidance, and driving directions. And

---

<sup>29</sup> Warnock, *The Object of Morality* (n 26) 64-65.

<sup>30</sup> Warnock, *The Object of Morality* (n 26) 65.

they have a special name: ‘rules of thumb’.<sup>31</sup> The key point about a rule of thumb is, as Frederick Schauer says, that it ‘establishes a confidence level necessary for taking actions inconsistent with the rule’.<sup>32</sup> ‘If we believe that a paragraph will better fit the flow of an argument if it discards the usual initial topic sentence, or that this particular type of car will start more quickly with one rather than three pumps of the accelerator, our allegiance to the rule comes to an end’,<sup>33</sup> which ‘is because these are rules of thumb, providing useful guides for the routine case’.<sup>34</sup>

The confidence level associated with a rule of thumb can vary considerably. Schauer uses the example of Captain Prescott’s famous direction, ‘Don’t fire until you see the whites of their eyes!’, and he supposes that it becomes a rule of thumb among the Captain’s riflemen.

Suppose that the whites-of-their-eyes rule has *ceteris paribus* status. Riflemen are expected to make their own calculations with respect to the optimal shooting moment, but should rely on the whites-of-their-eyes rule when, all things considered, their own calculations provide no answer.<sup>35</sup>

If a rifleman does not believe the balance of reasons favours one moment over another, having the rule will lead him to act as if he believes he should fire at the whites-of-their-eyes moment. If he believes another moment is best, though, the rule will not lead him to act that way.

Suppose instead that the whites-of-their-eyes rule has not merely *ceteris paribus* status, but is instead formulated in such a way that riflemen should fire at some moment other than the whites-of-their-eyes moment if and only if they are convinced that the whites-of-their-eyes moment is not the best moment for firing. .... If the rifleman, despite not having the level of confidence

---

<sup>31</sup> There is a great deal of terminological disagreement about “rule of thumb”. Raz uses the term (*Practical Reason and Norms* (n 3) 59-62) somewhat idiosyncratically to describe a type of (what I am calling) strict rule, which is distinguished by its justification as a time- or labour-saving device. Others use the term to mean something like a reminder of a reason. See, for example: Goldman (n 2) 14-19; Shapiro and McClennen (n 26) 363-364. My usage is largely consistent with Schauer’s. He defends his choice in *Playing by the Rules* (n 12) 105 n 36.

<sup>32</sup> Schauer, *Playing by the Rules* (n 12) 108.

<sup>33</sup> Schauer, *Playing by the Rules* (n 12) 4.

<sup>34</sup> Schauer, *Playing by the Rules* (n 12) 4. Emphasis omitted.

<sup>35</sup> Schauer, *Playing by the Rules* (n 12) 108.

established by the rule, adheres to his own justification-based judgment, he has once again broken the rule. And the same analysis would apply if the rule of thumb were formulated in terms of certainty instead of conviction.<sup>36</sup>

If a rifleman is not ‘convinced’ or ‘certain’ that the overall balance of reasons favours a moment other than the whites-of-their-eyes moment, the rule will lead him to act as if he believes he should fire at the whites-of-their-eyes moment. If he believes another moment is best, he will not be so disposed.

### *Error-Avoidance and Strictness*

Earlier I described what I termed ‘ordinary decision-making’. It starts with the balancing process, leads to a judgment of what ought to be done, and culminates in a decision. One of the common problems with ordinary decision-making is that it can be corrupted by incorrect beliefs about what reasons there are and how they compare by weight. When a person decides what to do in the ordinary way, either type of error can lead to an incorrect judgment of what ought to be done, and ultimately a wrong action. A traditional justification for having a rule is to compensate for these errors.

Suppose you live somewhere with deadly black widow spiders. Since one bite from a black widow can kill you, you ought to avoid them. Almost all black widows have red markings on their abdomens, so you can be reasonably confident that a spider *without* red markings is not a black widow. The converse is not true, however, because there is another type of spider, known as a ‘false’ black widow, which also has red markings on its abdomen but is harmless. There are visible differences between real and false black widows, and you can distinguish the two, more or less. Given the risks, though, more or less is not good

---

<sup>36</sup> Schauer, *Playing by the Rules* (n 12) 108.

enough. You need to make sure you do not act on your judgment of whether a spider is a real black widow or a false black widow.

Having a strict rule to avoid spiders with red markings can help you to accomplish that goal. Having this rule means that you will act as if you believe that *every* spider with red markings ought to be avoided, not just the ones that you believe are black widows. Without your rule, you might err by (1) avoiding a false black widow or (2) approaching a real black widow. With your rule, your errors are much more likely to be of the first type. You might end up erring more often, if you encounter more false black widows than real black widows. Within limits, though, having the rule is still justified, because the costs of the two types of errors are so unequal: death, potentially, versus minor inconvenience.

For experts, having a strict rule to avoid spiders with red markings may be unnecessary. Suppose you are sufficiently knowledgeable that, if you are *very* confident that a particular spider with red markings is *not* a real black widow, then you are certain to be correct. You are never wrong when you are that confident. In your case, the rule to avoid spiders with red markings should be a rule of thumb, not a strict rule. It will lead you to act as if you believe spiders with red markings ought to be avoided, except when you are very confident that *this* spider need not be avoided. Your expertise gives you an advantage over non-experts. You can afford to have a rule that leads to less inconvenience but no additional risk.

### *Strength and Strictness Combined*

So far I have discussed a dimension of strength and a dimension of strictness. These dimensions intersect, generating four types of rules. For example, if a society has a rule that  $\mathcal{A}$ s ought to  $\phi$ , and that rule is strong as well as strict, then the members of that society act as

if they believe  $\mathcal{A}$ s ought to  $\phi$ , all things considered, regardless whether they disbelieve that an  $\mathcal{A}$  ought to  $\phi$ . If, on the other hand, a society has a weak rule of thumb that  $\mathcal{A}$ s ought to  $\phi$ , then its members act as if they believe  $\mathcal{A}$ s ought to  $\phi$ , relative to some type of reason, but not when they disbelieve that an  $\mathcal{A}$  ought to  $\phi$ , relative to that type of reason. The remaining two categories – strong rules of thumb, and rules that are weak but strict – can be thought of similarly. And the same conditions apply, appropriately modified, to personal rules.

#### 5.4 CONCLUSION

At the end of the last chapter I concluded that for a society to have a rule is for its members to have a certain attitude. In this chapter I have tried to find out more about that attitude. It leads people to act as if they believe that rule ought to be complied with, all things considered, or relative to a type of reason. It does not entail the belief that the balance of reasons always favours compliance with that rule. And, sometimes, the motivational force of the attitude is lacking when you are confident to some degree that, in a particular case, the balance of reasons does not favour compliance with that rule. Based on these points, I identified dimensions of strength and strictness along which rules may range.

A question I raised but did not answer was whether the attitude in question entails the belief that the rule ought to be complied with, in all-things-considered or relative sense, as appropriate. Were the balance of reasons to determine what ought to be done, it would be clear that having a rule does *not* entail the belief that the rule ought to be complied with. But, according to Raz, the balance of reasons may not determine what ought to be done, because exclusionary reasons defeat the reasons they exclude without outweighing them. The next chapter discusses exclusionary reasons and what significance they may have for rule-possession.

# 6 Protected Reasons and Rule-Possession

---

## 6.1 INTRODUCTION

The last two chapters have been an information-gathering exercise meant to identify some of the features of the attitude on which rule-possession depends. One attitude that seems to have the right features is a belief in a protected reason to comply with a rule. In this chapter I investigate whether for a society to have a rule simply means that its members have such a belief. I argue that it does not, but that the reasons why are instructive. This chapter advances my overall project in part by showing why one account of what it is to have a rule should be rejected. That is worth showing so long as this account is plausible on its face, which I argue it is. This chapter also narrows my search by revealing more about rule-possession and how it is related to the belief that the rule ought to be complied with.

Let me be as clear as possible about one thing. In this chapter I consider a theory of rule-possession that builds on some of Raz's claims about protected reasons. I try to undermine that theory by making certain limited points against Raz's claims about protected reasons. However, and this is crucial, I do *not* attribute this theory of rule-possession to Raz and, in fact, I do not think he favours it or one like it. Indeed, I do not attribute *any* theory of what it is to have a rule to Raz. He may not even have a view on that topic, and it would not matter for present purposes if he did.

## 6.2 PROTECTED REASONS AND RULES

### *Two Approaches*

When a society has a rule, its members act as if they believe that rule ought to be complied with. This is one of the points I made in the last chapter. (We can set aside for the moment whether the ‘ought’ is all-things-considered or relative.) Now, it is natural to think that, if you act as if you believe something ought to be done, it is because you believe that is what ought to be done. The temptation is to say:

- (1) If a society has a rule that  $\mathcal{A}$ s ought to  $\phi$ , then its members believe that  $\mathcal{A}$ s ought to  $\phi$ .

Another point from the last chapter is that, sometimes, a society has a rule but its members do not believe that the balance of reasons favours complying with that rule. A second natural thought is that what ought to be done is just what the balance of reasons favours. The temptation is then to say:

- (2) Sometimes, a society has a rule that  $\mathcal{A}$ s ought to  $\phi$ , but its members do *not* believe that  $\mathcal{A}$ s ought to  $\phi$ .

Claims (1) and (2) are incompatible, so we need to avoid committing ourselves to both of them. There are two approaches. The first approach is to accept (1) and reject (2). Provided that people who have a rule are not often and seriously conceptually confused, rejecting (2) means breaking the connection between ‘ought’ and the balance of reasons. The second approach is to reject (1) and accept (2). That requires an explanation of why people act as if they believe what really they do not.

Here is an example that illustrates the two approaches. Judges have a rule that they are to apply the law and that leads them to act as if they believe they ought to do so. The obvious explanation for why they behave this way is that they believe they ought to apply the law. It is nothing more mysterious than that. However, it seems that judges do not believe the balance of reasons always favours applying the law. Recall that this was what Hart seemed to suggest. That is problematic because what ought to be done is naturally thought to be determined by the balance of reasons. So, the ‘obvious’ explanation or the ‘natural’ thought must be incorrect. Either judges do not really believe they ought to apply the law, though they act as if they do, or the reasons for and against applying the law do not determine whether it ought to be applied.

In later chapters I argue for a version of the second approach. I argue, in other words, that people who have a rule may not believe it ought to be complied with, despite acting as if they do. For now, I want to consider a version of the first approach.

### *Motivating a Theory*

The key to the first approach is to give a plausible explanation of why what ought to be done may differ from what the balance of reasons favours. Raz, of course, claims that the two standards may diverge because an exclusionary reason prevails over the reasons it excludes without outweighing them. If that claim is plausible, it could give us a compelling, belief-based theory of what it is to have a rule. Let us suppose that this claim is plausible and that exclusionary reasons do take priority over excluded reasons. If there is a protected reason for  $\mathcal{A}$ s to  $\phi$ , then there is a reason for  $\mathcal{A}$ s to  $\phi$  and a reason that excludes at least some conflicting reasons. The excluded reasons are defeated, so  $\mathcal{A}$ s ought to  $\phi$ , relative to all these reasons. The excluded reasons may not be outweighed, so the balance of all these reasons may not

favour  $\mathcal{A}$ s  $\phi$ ing. Supposing now that people *believe* there is a protected reason for  $\mathcal{A}$ s to  $\phi$ , it is easy to see why they might also believe that  $\mathcal{A}$ s ought to  $\phi$ , and so act *as if* they believe the same, even though they do not believe the balance of reasons favours  $\mathcal{A}$ s  $\phi$ ing.<sup>1</sup>

At this point, a protected reasons-based theory of having a rule begins to look very attractive. That theory would say, essentially, that some person or some society has a rule that  $\mathcal{A}$ s ought to  $\phi$  just if that person or enough members of that society believe there is a protected reason for  $\mathcal{A}$ s to  $\phi$ . Two of the advantages of this theory are suggested in the last paragraph. The theory explains why having a rule leads people to act as if they believe it ought to be complied with – it is because they really believe that. It explains why having a rule does *not* entail believing that the balance of reasons favours compliance with that rule – it is because a protected reason may not be the weightiest reason. And it turns out that the theory is safe from Warnock’s Objection. Believing that  $\mathcal{A}$ s ought to  $\phi$  does not entail believing in a protected reason for  $\mathcal{A}$ s to  $\phi$ , which, according to this theory, is why the members of a society may believe that  $\mathcal{A}$ s ought to  $\phi$  without having a rule that  $\mathcal{A}$ s ought to  $\phi$ .

Its advantages mean that we need to take the protected reasons-based theory seriously. Taking it seriously means doing two things. It means, first, seeing whether this theory can account for what else we know about having a rule and, in particular, the way in which having rules of varying strengths and strictnesses lead people to act. That is what I do in the rest of this section. It also means thinking carefully about whether there is a good basis for saying that exclusionary reasons take priority over excluded reasons. That is what I do in the next section.

---

<sup>1</sup> This is the result whether people believe there is a protected reason for  $\mathcal{A}$ s to  $\phi$  – meaning, a fact that is at once both a reason for  $\mathcal{A}$ s to  $\phi$  and a reason that excludes some conflicting reasons – or merely believe there is a reason for  $\mathcal{A}$ s to  $\phi$  and a reason that excludes some conflicting reasons. I am opting for the first alternative because it is simpler to say. A full theory of rule-possession would resolve this issue one way or the other.

Let me add that neither this theory nor any theory of rule-possession is endorsed or considered by Raz. It is a theory I have formulated, for the sake of argument, and I have drawn on Raz's claims about exclusionary reasons in order to do so. This is also a good time to point out that the relevance of much of the material from Chapter 2, on reasons, should now be clear. I have referred to reasons repeatedly in the last two chapters and in the last chapter I relied on the idea of one reason 'defeating' another reason and on the distinction between all-things-considered and relative 'oughts'. As I have just explained, and will now show in more detail, exclusionary and protected reasons may be particularly relevant to rule-possession.

### *Strong Rules and Weak Rules*

One of the virtues of the protected reasons-based theory is that it can easily account for the difference between strong and weak rules. Remember that having a strong rule leads people to act as if they believe the rule ought to be complied with, all things considered. When people have a weak rule, things are the same, except that the 'ought' is relative to a type of reason. If we think of rule-possession as based on a belief in a protected reason, then the scope of the protected reason accounts for the strength of the rule.

Consider the constitutional convention that the Queen ought to assent to legislation. The convention is (or we can assume it is) a strong rule, held among constitutional actors. They have an attitude that leads them to act as if they believe the Queen ought to assent to legislation, all things considered. Now, suppose that constitutional actors believe there is a *fully* protected reason for the Queen to assent to legislation – that is to say, a fact that is a reason for her to grant assent and a reason that excludes *all* conflicting reasons. They will believe that all reasons against the Queen granting assent are excluded and so defeated,

leaving unchallenged the reason that favours her granting assent. They will believe the Queen ought to assent to legislation, even if they do not believe the reasons for her to do so are weightier than the opposing reasons.<sup>2</sup> So, in these respects, a society that has a strong rule is the same as a society whose members believe in a fully protected reason to comply with that rule.

What about weak rules? Consider the Kashruth rule that prohibits Jews from eating pork. On Schauer's understanding, this rule is weak in the sense that it is 'overrideable' by reasons of survival, though not by less important reasons. Assuming Schauer's understanding is correct, those who have this rule act as if they believe that, relative to non-survival-related reasons, no one ought to eat pork. Again, a belief in a type of protected reason leads people to act in just the same way. If people believe there is a reason not to eat pork that is protected from all but survival-related reasons, then they should also believe that pork ought not to be eaten, relative to non-survival-related reasons. That would be true even if they had no belief about whether the balance of those reasons favoured eating pork or not. In these crucial ways, a society that has a weak rule is the same as a society whose members believe in a partially protected reason to comply with that rule.

### *Strict Rules and Rules of Thumb*

Thinking of rule-possession in terms of a belief in a protected reason provides a straightforward account of strict rules. However, it makes it difficult to account for rules of thumb. Recall that the difference between possessing the two types of rules is marked by what happens when people believe that the balance of reasons does *not* favour compliance

---

<sup>2</sup> I am assuming, here and throughout this section, that no relevant reasons are cancelled, and that there are no other relevant second-order reasons.

with the rule. If the rule is strict, people act as if they believe it ought to be complied with anyway; not so if the rule is a rule of thumb.

Suppose that the convention that the Queen ought to assent to legislation is a strict rule. Constitutional actors proceed as though they believe the Queen ought to assent to legislation, even when they believe (to any degree of confidence) that the balance of reasons does not favour her granting assent. That is easily explained if the actors believe in a protected reason for the Queen to assent to legislation. They will believe there is a reason for the Queen to assent to legislation that remains undefeated, no matter how clearly outweighed. They will believe, as a result, that the Queen ought to assent to legislation, even when they believe balance of reasons tilts the other way. And so they will act as if they believed the same. (This example is of a rule that is both strict and strong, but something similar could be said of rules that are serious and weak.)

What is an advantage when it comes to strict rules is a disadvantage with respect to rules of thumb. When people believe there is a protected reason for an action, they act as if they believe the action ought to be performed, *even when* they believe the balance of reasons does not favour the action. That is what possession of a serious rule and belief in a protected reason have in common. But when people have a rule of thumb, they act as if they believe it ought to be complied with, *except when* they believe (to some degree of confidence) that the balance of reasons does not favour the action. That is how possession of a rule of thumb and belief in a protected reason differ.

Recall Schauer's example of the riflemen from the last chapter. Schauer imagines the riflemen have a rule of thumb that tells them not to fire until they 'see the whites of their eyes'. Having the rule leads the riflemen to act as if they believe they ought to fire at the whites-of-their-eyes moment, except when they are sufficiently confident that there is more merit in firing at another moment. In contrast, believing in a protected reason to fire at the

whites-of-their-eyes moment would lead the riflemen to treat the merits of firing at various moments as irrelevant, no matter how clearly they favoured firing at a particular moment. (It does not matter whether the rule of thumb is strong or weak, or whether the reason is fully or partially protected.)<sup>3</sup>

### *A Preliminary Assessment*

If protected reasons take priority over ordinary reasons, then there is much to be said for a protected reasons-based theory of rule-possession. That theory explains the way in which people act when they have a rule. It explains the relation between having a rule and key beliefs. It explains the difference between strong and weak rules, and it explains strict rules. The only thing it seems incapable of explaining is rules of thumb. How serious a problem is that? Some scholars do not think of rules of thumb as ‘real’ rules.<sup>4</sup> Presumably they would see it as an advantage of the protected reasons-based account that it only addresses strict rules. Even scholars more accepting of rules of thumb might not be too troubled. True, they might say, other things being equal, a unified account would be preferable, but this account still has a great deal to be said for it. Furthermore, it is not clear that there is a unified account on offer. Perhaps we ought to adopt the protected reasons-based account as far as it goes and then seek to explain rules of thumb in some other way.

---

<sup>3</sup> Could we fix the problem by making the protected reason conditional? A rifleman might have a belief with the following content: *unless I am sufficiently confident that the balance of reasons favours firing at a different moment, there is a protected reason to fire at the whites-of-their-eyes moment.* The rifleman would then act in at least roughly the same way as he would be were to have a rule of thumb. The problem with the proposal is that it saddles the rifleman – and, indeed, anyone with a rule of thumb – with an extremely odd belief. The rifleman would have to believe that whether he ought to fire at the whites-of-their-eyes moment turned on his belief about the merits of firing at that moment and not, as seems much more likely, on the actual merits of firing at that moment. Such beliefs are too strange and difficult to maintain for us to attribute them to every person who has a rule of thumb.

<sup>4</sup> Alan Goldman, for example, classifies rules into ‘genuine’ rules and rules of thumb: *Practical Rules: When We Need Them and When We Don’t* (Cambridge University Press, 2002) 2-3, 26-17.

Given that the literature emphasizes strict rules, that there is no unified account yet on offer, and that I may have made a mistake, I will not treat the failure of the protected reasons-based theory to account for rules of thumb as a fatal flaw. That means I am not in a position to reject that theory at this stage in the argument. It also makes it more important to test the assumption I have been making in this section, and on which this theory depends, namely, that exclusionary reasons take priority over the reasons they exclude.

### 6.3 RULES AND PARTIAL CONFLICTS

At the beginning of this chapter, I said that what we know about having a rule leaves us with two options. Either a society may have a rule even though its members do not believe it ought to be complied with, or what ought to be done may diverge from what the balance of reasons favours. I said, too, that the viability of the second approach depends on there being a persuasive argument for why ‘ought’ and the balance of reasons may diverge. We need such an argument because the intuitive view, as well as the standard or conventional view among moral philosophers, is to the contrary.<sup>5</sup>

In the last section I assumed that what ought to be done may not be what the balance of reasons favours, because I assumed that exclusionary reasons prevail over the reasons they

---

<sup>5</sup> Raz acknowledged that his view was unconventional. He said that ‘[t]he distinction between first-order and second-order reasons for action has not been recognized or discussed by philosophers’. He went on to say:

According to our intuitive conception of practical conflicts such conflicts are to be resolved by assessing the relative strength or weight of the conflicting reasons and determining what ought to be done on the balance of reasons. To put it another way, one ought always to do whatever one has a conclusive reason for doing. Or, which is another way of saying the same, one ought always to act on the balance of reasons.

Raz expressed this ‘intuitive conception of conflict resolution’ in the form of the following principle: ‘It is always the case that one ought, all things considered, to do whatever one ought to do on the balance of reasons’. Raz, *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990) 36. For a small sample of moral philosophers who have recently adopted or proceeded on the basis of this ‘intuitive’ view about the balance of reasons and ‘ought’ (in the relevant sense), see: Judith Jarvis Thomson, *Goodness and Advice* (Amy Gutman ed, Princeton University Press 2001) 74-76; Bart Streumer, ‘Can Consequentialism Cover Everything?’ (2003) 15 *Utilitas* 237, 243; Ingmar Persson, ‘A Consequentialist Distinction between What we Ought to Do and Ought to Try’ (2008) 20 *Utilitas* 348, 354; Stephen Finlay, ‘What *Ought* Probably Means, and Why You Can’t Deatch It’ (2010) 177 *Synthese* 67; Benjamin Kiesewetter, ‘“Ought” and the Perspective of the Agent’ (2011) 5 *Journal of Ethics & Social Philosophy* 1.

exclude, regardless of weight. I showed that, granting those assumptions, a protected reasons-based account can explain much of what we know about rule-possession. Now I want to drop these assumptions. I want to consider whether there is a good argument to the effect that exclusionary reasons work the way Raz says they do, and thus whether Raz has given us grounds for thinking that ‘ought’ and the balance of reasons may diverge. Let me emphasize the modesty of my aim. I do not try to show that it is *false* that exclusionary reasons prevail over the reasons they exclude, whatever their weight. That is a more difficult task than the one I attempt. Rather, I claim that Raz has not given us reason to believe it is *true* that exclusionary reasons work as he supposes they do. In the absence of a reason for accepting Raz’s view, and given its unorthodoxy, we are justified in withholding our acceptance.

The majority of this section is a fairly detailed discussion of one way in which reasons may conflict. I take this methodical approach because it seems the most careful, thorough, and I think the fairest way to present and assess Raz’s views about exclusionary reasons.

### *Raz’s Claim*

Why does Raz think that an exclusionary reason always prevails over a reason it excludes? In the first edition of *Practical Reason and Norms*, where he introduced exclusionary reasons, Raz did not say. Nor did he say how exclusionary and excluded reasons conflict; he simply said they do not strictly conflict.<sup>6</sup> He was more explicit in the postscript to the second edition. There, he said that exclusionary reasons ‘partially conflict’ with excluded reasons, and that this explains why exclusionary reasons prevail.

---

<sup>6</sup> Raz, *Practical Reason and Norms* (n 5) 40.

In the book I suggested that exclusionary reasons do not compete in weight with the reasons they exclude; rather, they always win in such conflicts. It is possible to account for this result by regarding the relations between exclusionary reasons and the excluded reasons as governed by general considerations affecting partial conflicts. Exclusionary reasons partially conflict with the reasons they exclude.<sup>7</sup>

Raz says here that exclusionary reasons prevail over excluded reasons because of ‘general considerations affecting partial conflicts’. To assess that claim, we need to know more about partial conflicts, how they are resolved, and what relevance these issues have for exclusionary reasons.

### *Types of Partial Conflicts*

Raz does not say very much about partial conflicts. In an early article titled ‘Reasons, Requirements, and Practical Conflicts’, he says:

If an agent has reasons to do [ $\phi$ ] and reasons to do [ $\psi$ ] and if his doing one of those acts will affect any aspect of his ability to perform the other act then he is confronted with a practical problem of (at least) a partial conflict.<sup>8</sup>

Based on this description, Raz identifies several types of partial conflict, only one of which is relevant here.<sup>9</sup> Reasons to  $\phi$  and  $\psi$  partially conflict if ‘some ways of  $\phi$ -ing are compatible with some ways of [ $\psi$ ]-ing, but certain ways of  $\phi$ -ing are incompatible with some ways of [ $\psi$ ]-ing’.<sup>10</sup> Let us call this ‘type 1’ partial conflict. Here is an example: if there is a blue circle and a blue square, and a red circle and a red square, then there is a type 1 partial conflict between a

---

<sup>7</sup> Raz, *Practical Reason and Norms* (n 5) 189.

<sup>8</sup> Joseph Raz, ‘Reasons, Requirements, and Practical Conflicts’ in Stephen Korner (ed), *Practical Reason* (Yale University Press 1974) 26. Raz summarizes his discussion of partial conflicts in *Practical Reason and Norms* (n 5) 202 n 4.

<sup>9</sup> There are two other types of partial conflicts he mentions: (1) Reasons partially conflict when they ‘are reasons to achieve two goals and where the full realization of one goal though compatible with the realization of the second goal to a certain extent is incompatible with its realization in the highest degree’; (2) Reasons to  $\phi$  and  $\psi$  partially conflict if ‘ $\phi$ -ing is compatible with ( $\psi$ )-ing on most occasions but incompatible with it in some contexts’. Raz, *Practical Reason and Norms* (n 5) 202 n 4; see also, Raz, ‘Reasons, Requirements, and Practical Conflicts’ (n 8) 25-26.

<sup>10</sup> Raz, *Practical Reason and Norms* (n 5) 202 n 4.

reason to pick a blue shape and a reason to pick a circle, because there is a way to satisfy both reasons (picking a blue circle) and a way to satisfy the first but not the second (picking a blue square).

Reasons that partially conflict in the type 1 sense do *not* strictly conflict.<sup>11</sup> They do not strictly conflict because, by definition, it is possible to satisfy them both, whereas reasons that strictly conflict are reasons for actions that are incompatible. However, because a reason for an action entails a reason for a specific way of performing that action<sup>12</sup>, it *is* possible to frame type 1 partial conflicts in terms of a strict conflict – just not a strict conflict between the reasons that partially conflict. These are the relevant reasons in the example:

1. Reason to pick a blue shape.
  - a. Reason to pick a blue square.
  - b. Reason to pick a blue circle.
2. Reason to pick a circle.
  - a. Reason to pick a red circle.
  - b. Reason to pick a blue circle.

Some of the reasons entailed by (1), the reason to pick a blue shape, strictly conflict with some of the reasons entailed by (2), the reason to pick a circle. Specifically, (1a) strictly conflicts with each of (2a) and (2b). Other reasons entailed by (1) and (2) do not strictly conflict. Specifically, (1b) does not strictly conflict with (2b). In general, reasons partially

---

<sup>11</sup> Recall Raz's definition of strict conflict: ' $p$  strictly conflicts with  $q$  relative to  $x$  and  $\phi$  if, and only if,  $R(\phi)p,x$  and  $R(\neg\phi)q,x$ , i.e. that  $p$  is a reason for  $x$  to  $\phi$  and that  $q$  is a reason to refrain from  $\phi$ -ing', where ' $R(\phi)p,x$ ' means ' $p$  is a reason for an agent  $x$  to  $\phi$ '. Raz, *Practical Reason and Norms* (n 5) 25.

<sup>12</sup> 'One has reason to do whatever will facilitate conformity with reason. More narrowly, one has reason to do whatever is a sufficient condition for the realisation of some good, including the good of conformity with reason'. Raz, *Practical Reason and Norms* (n 5) 182.

conflict in the type 1 sense if some of the reasons they entail strictly conflict and others do not.<sup>13</sup>

Raz returned to partial conflicts in 'Facing Up' and the postscript to *Practical Reason and Norms*. In the postscript, he gives his definition of strict conflict, and then says:

Sometimes if a reason is overridden it is completely frustrated: that is, the good that conforming with it would have secured will not be achieved. But often this is not the case. Often the good that conformity with the reason would achieve can be secured, completely or partly, in some other way, which may be open to the agent or to some other people. ... Whenever at least one of two conflicting reasons need not be completely frustrated, the conflicts are partial conflicts.<sup>14</sup>

The idea seems to be the following. Sometimes reasons strictly conflict<sup>15</sup> but it is possible to realize the values that conformity with each would have realized, in which case the reasons also partially conflict. In other words, reasons that strictly conflict also partially conflict if it is possible to satisfy the operative parts of each reason. Call this 'type 2' partial conflict.

The example above includes a type 2 partial conflict as well. The reason to pick a blue square, (1a), strictly conflicts with the reason to pick a circle, (2), but the reasons also partially conflict, because picking a blue circle secures the value underlying the reason to pick a blue shape and the value underlying the reason to pick a circle (assuming these are different values). By the same reasoning, the reason to pick a red circle, (2a), partially conflicts with the reason to pick a blue shape, (1).

Raz thinks exclusionary and excluded reasons partially conflict. He does not say how they partially conflict, but it must be a type 1 partial conflict. For suppose  $p$  is a reason to  $\phi$

---

<sup>13</sup> Raz, 'Reasons, Requirements, and Partial Conflicts' (n 8) 26: "To the extent that a reason for an act is also a reason for any act by which it can be done, [this] type of partial conflict can also be characterized in terms of diametrical [i.e., strict] conflict: A reason for  $A$  conflicts with a reason for  $B$  if  $A$  can be done by  $A'$  and  $B$  by  $B'$  and the reasons for  $A'$  and for  $B'$  conflict diametrically [strictly]."

<sup>14</sup> Raz, *Practical Reason and Norms* (n 5) 188.

<sup>15</sup> The reasons must strictly conflict because in the quoted passage Raz says one reason could 'override' another and only strictly conflicting reasons are capable of overriding each other: Raz, *Practical Reason and Norms* (n 5) 26.

and  $e$  is an exclusionary reason not to  $\phi$  for  $p$ . There are two ways to satisfy  $p$ : you could  $\phi$  for  $p$  or you could  $\phi$  not for  $p$ .<sup>16</sup> There are also two ways to satisfy  $e$ : you could not  $\phi$  or you could  $\phi$  not for  $p$ . These are the relevant reasons:

3. ( $p$ ) Reason to  $\phi$ .
  - a. Reason to  $\phi$  for  $p$ .
  - b. Reason to  $\phi$  not for  $p$ .
4. ( $e$ ) Reason not to  $\phi$  for  $p$ .
  - a. Reason not to  $\phi$ .
  - b. Reason to  $\phi$  not for  $p$ .

The structure of the example exactly mirrors the structure of the shapes example. Some of the ways of satisfying (3) are compatible with some of the ways of satisfying (4). Specifically, it is possible to satisfy both (3b) and (4b). Some of the ways of satisfying (3) are incompatible with some of the ways of satisfying (4). Specifically, it is not possible to satisfy (3a) and either (4a) and (4b). Accordingly, (3) and (4) – the exclusionary reason and the excluded reason – partially conflict, and the conflict is of type 1. There is *no* type 2 partial conflict between the exclusionary and excluded reason, in part because the reasons do not strictly conflict. (As in the shapes example, there may be a type 2 conflict between (4) and (3a) or between (3) and (4a).)

---

<sup>16</sup> Raz, *Practical Reason and Norms* (n 5) 182: ‘Complying with a reason [ie, acting for a reason] will, trivially, lead to, or rather constitute, conformity with it. Hence one always has reason to comply’.

## *Resolution of Partial Conflicts*

I have set out the relevant types of partial conflict and I have said that conflicts between exclusionary and excluded reasons are type 1 partial conflicts. Let me now turn to the resolution of partial conflicts. In the postscript to *Practical Reason and Norms*, Raz says:

Whenever at least one of two conflicting reasons need not be completely frustrated, the conflicts are partial conflicts. Let us take the case of two conflicting reasons, one of which, if overridden, would be completely frustrated, whereas the other need not be frustrated at all (i.e. there is an equally effective alternative way of realising the good that conformity with it will have achieved). In such a case the reason which would be frustrated overrides the one that will not be. The overridden reason will then serve as a reason for an alternative action which will achieve the good that conformity with it would have secured.<sup>17</sup>

I understand Raz this way: supposing two reasons strictly conflict (they must strictly conflict because only reasons that strictly conflict may ‘override’ each other<sup>18</sup>) and it is possible to satisfy the first reason and satisfy the operative part of the second reason, but not vice versa, then the first reason overrides the second.

The principle Raz is claiming is a type of dominance principle: you ought not to do what will realize just one value if you could do otherwise and realize a second value as well. Or, negatively, you should not frustrate a value unnecessarily. Expressed this second way, Raz is essentially claiming parsimoniousness as a principle of conflict resolution. He allows that the principle will have to be modified to account for special cases.<sup>19</sup> In its broad outline, then, what I term the ‘parsimony principle’ says that, for any reasons  $p$  and  $q$ , if these reasons

---

<sup>17</sup> Raz, *Practical Reason and Norms* (n 5) 188.

<sup>18</sup> Raz, *Practical Reason and Norms* (n 5) 26.

<sup>19</sup> He mentions three cases in *Practical Reason and Norms* (n 5) 188-189. First, the ‘alternative action’ may be unlikely to be performed. For example, ‘it might be that only someone else can take it, and that person is unlikely to be appropriately motivated to do so’. Second, ‘the principle has to be refined to take account of the fact that the reason for the alternative action itself may conflict with yet another reason’. Third, there are cases ‘when the reason which is not completely frustrated is partially frustrated’, and in such cases ‘depending on its importance and on the degree that it will be frustrated, it may override’. These qualifications only strengthen my ultimate criticism of Raz’s claim with respect to exclusionary reasons, so I shall set them aside. See also: Raz, ‘Facing Up: A Reply’ (1989) 62 *Southern California Law Review* 1153.

strictly conflict, it is possible to satisfy  $p$  and realize the operative part of  $q$ , and it is not possible to satisfy  $q$  and realize the operative part of  $p$ , then  $p$  overrides  $q$ . All the reasons to which the parsimony principle applies partially conflict in the type 2 sense. None of the reasons to which it applies partially conflict in the type 1 sense, because none of those reasons strictly conflict.

In ‘Facing Up’, Raz gives an example that can be used to illustrate the parsimony principle:

Assume that it is wrong to eat in a hurry. One should always eat slowly, in a measured way. This reason, one may say, partially conflicts with any other reason for eating. Certain ways of eating, certain ways of satisfying the reasons for eating, violate it. But other ways of eating are untouched. In such partial conflicts the question of which is the more important reason does not arise. Since one can satisfy both, one should always do so.<sup>20</sup>

There is a reason to eat and a reason to eat slowly<sup>21</sup>. The reason to eat entails a reason for each way of eating, including eating slowly and eating quickly. These are the relevant reasons:

5. Reason to eat.
  - a. Reason to eat quickly.
  - b. Reason to eat slowly.
6. Reason to eat slowly.

The reason to eat quickly, (5a), strictly conflicts with the reason to eat slowly, (6). It is possible to satisfy (6) and realize the value underlying (5a) – indeed, satisfying (6) ensures that – but (we can assume) it is not possible to satisfy (5a) and realize the value underlying (6). The

---

<sup>20</sup> Raz, ‘Facing Up’ (n 19) 1167.

<sup>21</sup> In the first sentence of the block quote, Raz asks us to assume that ‘it is wrong to eat in a hurry’, which would seem to mean assuming there is a reason not to eat in a hurry. Then he says ‘[o]ne should always eat slowly’, which would mean there is a reason to eat slowly. These are reasons for different things, because it is possible not to eat in a hurry by not eating, but it is not possible to eat slowly by not eating. I take it the second option is what Raz has in mind, both because of the overall context, and because of the reference to ‘any other reason for eating’ later in the same passage. This is also the way William Edmundson reads this passage: ‘Rethinking Exclusionary Reasons’ (1993) 12 *Law and Philosophy* 329, 338. If Raz has the first option in mind, then the example more closely parallels his example of exclusionary reasons (see text at n 23), but nothing substantive changes.

reasons stated by (5a) and (6) partially conflict, in the type 2 sense. According to the parsimony principle, (6) overrides (5a). Relative to these two reasons, one ought to eat slowly.<sup>22</sup>

### *Relevance to Exclusionary Reasons*

Raz presents his claim about the resolution of partial conflicts as relevant to exclusionary reasons. Immediately after introducing his eating example, Raz says:

In a similar way exclusionary reasons generate partial conflicts. The reason, if there is one, not to eat for pleasure partially conflicts with (first-order) reasons for eating. One can satisfy both, for one can eat without eating for pleasure. (Notice that one can have pleasure in eating even when one does not eat for pleasure.) That is why an exclusionary reason always triumphs in such conflicts.<sup>23</sup>

Breaking it down as I did before, there is a reason not to eat for pleasure, which entails a reason not to eat and a reason to eat but not for pleasure. In addition, a reason to eat entails a reason to eat for pleasure and a reason to eat but not for pleasure. These are the relevant reasons:

7. Reason to eat.
  - a. Reason to eat for pleasure.
  - b. Reason to eat not for pleasure.
8. Reason not to eat for pleasure.
  - a. Reason not to eat.
  - b. Reason to eat not for pleasure.

---

<sup>22</sup> The postscript to *Practical Reason and Norms* was written after 'Facing Up' (see *Practical Reason and Norms* (n 5) 213), so I take it that Raz's statements about partial conflicts in the postscript are the better guide to his views on the topic. The eating example from 'Facing Up' is, as I have interpreted it, an early illustration of ideas developed more fully in the postscript. Interpreting the example this way gives Raz's views continuity, which is why I think of it as the charitable interpretation.

<sup>23</sup> Raz, 'Facing Up' (n 19) 1167.

Raz claims that the exclusionary reason ‘wins’ this conflict and that it ‘triumphs’.<sup>24</sup> I take that to mean (8) is supposed to defeat (7) and not be defeated by it.

But why? The *only* thing Raz says about the resolution of partial conflicts in general is what I have expressed as the parsimony principle. And that principle *only* applies to type 2 partial conflicts, whereas conflicts between exclusionary and excluded reasons are type 1 partial conflicts. The parsimony principle is therefore inapplicable. It does not support the conclusion that (8) defeats (7). If I have misinterpreted Raz and it is not a condition of the parsimony principle that reasons strictly conflict, that principle might *still* not apply, because it is possible that the values underlying (7) or (8) could be realized in ways other than by satisfying those reasons. Even if the parsimony principle *did* apply, it might not favour the exclusionary reason over the excluded reason, but rather the opposite. It would depend on what other ways there are to realize the values underlying these reasons. So there are at least several obstacles to the conclusion that this exclusionary reason ‘triumphs’ over this excluded reason, let alone that exclusionary reasons always triumph over excluded reasons. (Under certain conditions, the parsimony principle would lead to the conclusion that this exclusionary reason defeats one of the reasons entailed by this excluded reason. See the discussion of Colin’s case below.)

At this point I have set out nearly everything Raz has said about partial conflicts. I have not found a convincing argument for why ‘general considerations affecting partial conflicts’ determine that exclusionary reasons always defeat excluded reasons. Since this is Raz’s only explicit argument for that conclusion, it seems that we should resist drawing it ourselves. But before I go on I want to consider whether a more limited claim would be easier to demonstrate. There could be some looser sense in which exclusionary reasons

---

<sup>24</sup> Raz uses a variety of equivalent terms, e.g., ‘win’, ‘upper hand’, ‘priority of exclusionary over the excluded reasons’. See Raz, *Practical Reason and Norms* (n 5) 40, 189, 190; Raz, ‘Facing Up’ (n 19) 1167, 1168.

prevail over excluded reasons. To show what that alternative might be, I turn to Raz's final example.

Raz supposes that Colin has promised his wife that 'in all decisions affecting the education of his son he will act only for his son's interest and disregard all other reasons'.<sup>25</sup> Colin now has to decide whether to send his son to a public school, and among the relevant reasons is the fact that if he 'does so he will be unable to resign his job in order to write the book he so much wants to write'.<sup>26</sup> Raz then says:

Colin's promise not to act out of concern for his own interest (that his son go to a state school, enabling him to resign his job and write a book) conflicts with that interest of his. They conflict in that there is a way of conforming with the excluded reason (i.e. sending his son to a state school in order to be able to resign his job) which the exclusionary reason, i.e. his promise, is a reason against. Colin may, however, decide to send his son to a state school (and this suits his interests), while not acting out of regard for his interest.<sup>27</sup>

So far Raz has said that it is possible to satisfy both an exclusionary and excluded reason by performing the action that the excluded reason favours but not for that reason. Now Raz makes the important claim:

Since it is possible to satisfy both the excluded reason and the reason that excludes it, the conflict between them is merely partial. Given that the conflict is only partial, and that complying with the excluded reason (Colin's sending his son to a state school in part because it is in his – Colin's – interest to do so) will completely frustrate the exclusionary reason (the promise) whereas not complying with the excluded reason will conform to the exclusionary reason while leaving a possibility that the agent will conform with the excluded reason, it follows that exclusionary reasons defeat the excluded reasons to the extent that one may conform with the excluded reason only in ways which conform with the exclusionary reason as well.<sup>28</sup>

These are Colin's reasons:

9. (Promise) Reason not to act for Colin's interests.

---

<sup>25</sup> The example is introduced in the first edition: Raz, *Practical Reason and Norms* (n 5) 39.

<sup>26</sup> Raz, *Practical Reason and Norms* (n 5) 39.

<sup>27</sup> Raz, *Practical Reason and Norms* (n 5) 189.

<sup>28</sup> Raz, *Practical Reason and Norms* (n 5) 189.

- a. Reason to send his son to public school.
  - b. Reason to send his son to state school not for Colin's interests.
10. (Interests) Reason to send his son to state school.
- a. Reason to send his son to state school for Colin's interests.
  - b. Reason to send his son to state school not for Colin's interests.

Raz says that by sending his son to state school because it is in his interests – i.e., by satisfying (10a) – Colin will ‘completely frustrate the exclusionary reason’, (9). On the other hand, by not satisfying (10a), Colin will satisfy the exclusionary reason ‘while leaving a possibility that [Colin] will conform with the excluded reason’, (10). From this, Raz seems to conclude that (9) prevails over (10a).

There are two points to note about the example and Raz's analysis of it. First, Raz assumes that if Colin sends his son to state school for his own interests it will ‘completely frustrate’ the exclusionary reason. Certainly that will mean Colin does not do as the exclusionary reason is a reason to do, but completely frustrating that reason is different. That a reason is completely frustrated means that ‘the good that conforming with it would have secured will not be achieved’.<sup>29</sup> It would seem, then, that Raz is assuming that conforming to the exclusionary reason is *necessary* to ‘secure the good that conforming with it would have secured’.

Let us accept this assumption for the sake of argument. In that case, with respect to (9) and (10a), all the conditions of the parsimony principle are satisfied. The two reasons strictly conflict, it is possible to satisfy (9) and realize the value underlying (10a), and we have assumed it is impossible to satisfy (10a) and realize the value underlying (9). It follows that (9) defeats (10a). That is all the argument shows. Assuming conformity to an exclusionary reason

---

<sup>29</sup> Raz, *Practical Reason and Norms* (n 5) 188.

is necessary to realize a certain value, and setting aside all the qualifications to the parsimony principle, an exclusionary reason defeats one of the reasons that are entailed by an excluded reason. Perhaps this is all Raz means to show. But, if so, it falls short of what he claimed on behalf of exclusionary reasons in his early work.<sup>30</sup> It does not follow that an exclusionary reason defeats an excluded reason, or that it takes ‘priority’ over it. More importantly, it does not follow that whether an action ought to be performed may be affected by the exclusion of reasons for or against it.

### *Significance for Rules*

Having set out and assessed Raz’s claims about partial conflicts and exclusionary reasons, let me now return to rules. I earlier presented a protected reasons-based theory of rule-possession. That theory promises to explain much, though not all, of what we know about rule-possession – but only if protected reasons prevail over other reasons regardless of weight. I argued in this section that we lack a convincing argument for that claim, which leaves the protected reasons-based theory without a solid foundation.

Someone might think that the protected reasons-based theory only depends on it being the case that what ought to be done may diverge from the balance of reasons. Although an exclusionary reason may not prevail over an excluded reason, according to the Parsimony Principle, under certain conditions, an exclusionary reason *does* prevail over one of the reasons entailed by the excluded reason and, more importantly, does so regardless of weight. Thus, exclusionary reasons, through the parsimony principle, sometimes circumvent the normal balancing process. This response retreats from the strong claim about exclusionary reasons,

---

<sup>30</sup> See n 47 and n 48 and accompanying text in Section 2.5.

while maintaining that belief in a protected reason can nonetheless explain rule-possession. But this response misunderstands what the parsimony principle does. It determines, in a narrow range of circumstances, whether one ought to do something in this way or that way, for this reason or that reason. Having a rule does something more. From the perspective of the person whose rule it is, the rule settles *what* ought to be done, not just *how* it ought to be done. The parsimony principle does not even apply in such circumstances. That is why, if the parsimony principle truly explains how exclusionary reasons work, then exclusionary reasons are poorly suited to explain how rules work.

Might it not be possible to argue for the priority of exclusionary reasons in some other way, not based on partial conflicts? I do not rule out the possibility. But I believe there is another theory of rule-possession for which there are compelling arguments, and I intend to set it out in the remaining chapters of this thesis. That theory could be modified if it turns out that exclusionary reasons do take priority over excluded reasons.<sup>31</sup>

## 6.4 CONCLUSION

I earlier said that having a rule leads people to act as if they believe that rule ought to be complied with. Given that people may have a rule but not believe that the balance of reasons favours complying with it, there were two options. The first option was that people who have a rule believe that it ought to be complied with, in which case what ought to be done must be capable of diverging from what the balance of reasons favours. The second option was that people who have a rule sometimes do not believe that it ought to be complied with, in which case some other attitude must be simulating belief.

---

<sup>31</sup> I explain what those changes would involve in Section 9.2.

In this chapter I proposed, for the sake of argument, a version of the first approach. The resulting theory, based on protected reasons, had many virtues. It could explain almost everything it needed to explain. It required, however, that protected reasons, and hence exclusionary reasons, take priority over at least some conflicting reasons. I argued that we have no basis for thinking that is the case. I did not try to show that the claim was false. But without an argument in its favour, we should not accept the claim as true. Since I am working largely within Raz's theory, I am left with override as the only way for one reason to defeat another reason. That means proceeding on the basis that the balance of reasons determines what ought to be done.

The second option is now forced on us. Having a rule leads people to act as if they believe that rule ought to be complied with, yet it neither entails nor is entailed by that belief. Some other attitude is at work, one which has the motivational force of belief, but which differs from it in other respects. To determine what attitude that is, I will look to the philosophy of action, in the next chapter, and the law of evidence, in Chapter 8. Once I have gathered the resources I need, I will return to rules in Chapter 9. In that chapter I also discuss the difference it makes if I have made a mistake and exclusion *is* a way for one reason to defeat another reason.

# 7 Acceptance and Counterfactuals

---

## 7.1 INTRODUCTION

People who have a rule act as if they believe the rule ought to be complied with, even though they may not believe that. What attitude leads them to act this way? An adequate theory of what it is to have a rule would answer this question, and in later chapters I propose a theory that does. This chapter lays the foundation for that theory.

I start in Section 2 by recasting the content of second-order reasons in counterfactual terms. That serves as a bridge to a range of other reasons that also have a counterfactual content. The reason I focus on is a reason to act in all ways as you would act were a certain event to have occurred, or what I call a reason to act ‘as if’. Some reasons to act ‘as if’ are reasons to act as if you believe something which actually you do not. When I return to rules, I argue that what is normally of value in having a rule is that it brings you to satisfy a reason to act as if you believe there is a certain reason for action you actually do not believe exists.

To explain how you can bring yourself to satisfy reasons to act ‘as if’, I turn in Section 3 to a distinction in the philosophy of action between belief, on the one hand, and acceptance, on the other. Both are propositional attitudes, and they have the same motivational force, but neither entails the other. Acceptance of a proposition leads you to act in the same ways as a belief in that proposition would. By accepting a proposition, you bring yourself to act as if you believe it. When I return to rules, I argue that what it is to have a rule is to accept that it ought to be complied with, whether or not you actually believe that.

## 7.2 COUNTERFACTUALS AND REASONS

### *Counterfactuals, Causation, and Second-Order Reasons*

Reasons to act and not act for other (valid) reasons are positive and negative second-order reasons. Exclusionary reasons are negative second-order reasons. Here is the definition I gave of ‘act for a reason’ in Section 2.4: that you  $\phi$  for  $p$  means you  $\phi$  because you believe that  $p$  and that  $p$  is a normative reason to  $\phi$ , and the explanatory connection works in the right way. In what follows, I will write ‘you believe there is a reason  $p$  to  $\phi$ ’ instead of ‘you believe that  $p$  and that  $p$  is a normative reason to  $\phi$ ’. Also, to simplify things, I will assume that whenever you  $\phi$  because you believe there is a reason  $p$  to  $\phi$ , the explanatory/causal works ‘in the right way’.

When do you  $\phi$  *because* you believe there is a reason  $p$  to  $\phi$ ? When does your believing that there is a reason  $p$  to  $\phi$  *cause* you to  $\phi$ ? My claim is that we can use a counterfactual test, at least as a rough guide.<sup>1</sup> Provided you actually  $\phi$ , and actually believe there is a reason  $p$  to  $\phi$ , your believing that caused your  $\phi$ ing if and only if you would not have  $\phi$ ed had you not believed there is a reason  $p$  to  $\phi$ . The motivation for this claim is simply that when we think about the causal effects of an event, like your believing there is a certain reason for a certain action, we tend to think about the difference that event’s occurrence made to what happened.

---

<sup>1</sup> David Lewis’s ‘Causation’ (1973) 70 *Journal of Philosophy* 556 sparked the recent interest in counterfactuals and causation. For helpful overviews of counterfactual theories of causation, see: John Collins, Ned Hall, and LA Paul, ‘Counterfactuals and Causation: History, Problems, and Prospects’ in their edited collection *Causation and Counterfactuals* (MIT Press 2004); Peter Menzies, ‘Counterfactual Theories of Causation’, *The Stanford Encyclopedia of Philosophy* (2008) <<http://plato.stanford.edu/entries/causation-counterfactual/>> accessed 8 July 2011.

The difference is measured by what happened that would not have happened absent the event.<sup>2</sup>

If you have a second-order reason to  $\phi$  for  $p$ , or to not  $\phi$  for  $p$ , then you actually believe there is a reason  $p$  to  $\phi$ . It is an open question, though, whether you actually  $\phi$ . There are two ways to satisfy an exclusionary reason not to  $\phi$  for  $p$ . You satisfy it either if you do not  $\phi$  or if you  $\phi$  but you would have done so anyway, i.e., had you not believed there is a reason  $p$  to  $\phi$ . Whichever is the case, your believing there is a reason  $p$  to  $\phi$  does not cause you to  $\phi$ . On the other hand, there is only one way to satisfy a positive second-order reason. You satisfy it only if you *both*  $\phi$  and would not have done so anyway. Expressed like this, second-order reasons have a counterfactual content.

The counterfactual test for causation reflects our commonsense understanding of causation in ordinary cases, but the two seem to diverge in more unusual cases. In particular, they may diverge in cases of *redundant causation*. These are cases in which there are two distinct events  $c$  and  $c'$ , and a third event  $e$ , and all three events actually occur. If either  $c$  or  $c'$  had occurred, in the absence of the other,  $e$  would have occurred; but if neither  $c$  nor  $c'$  had occurred,  $e$  would not have occurred.

There are different types of redundant causation cases. In *overdetermination* cases, the causal processes beginning from each of  $c$  and  $c'$  run to completion. Neither is preempted, and there is nothing to decide between  $c$  and  $c'$  as a cause of  $e$ . Here is an example of overdetermination by reasons:

Sam is offered a post in Middletown at an increased salary. Middletown is his wife's hometown and she wishes to return there. Both of these factors – the

---

<sup>2</sup> Lewis, 'Causation' (n 1) 557: 'We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it. Had it been absent, its effects — some of them, at least, and usually all — would have been absent as well.'

increased salary and his wife's wishes – may be reasons sufficient to induce Sam to accept the offer, in which case neither is a but-for reason.<sup>3</sup>

Sam would accept the offer had Middletown been his wife's hometown, or had the post come at a higher salary, in the absence of the other. But absent both, he would not accept the offer. Is each of Sam's reasons a cause of him accepting the offer? Is neither? Philosophers often doubt that there are clear intuitions about such cases. It is not obvious whether both conditions are causes or neither is. That seems to be true in Sam's case, too.

In *preemption* cases, it is clear that *c* or *c'* causes *e*, while the other does not. The first, the *preempted cause*, causes *e*; the second, the *preempted alternative*, does not cause *e*. There is a causal process that would have run from the preempted alternative to *e*, but that process was preempted – terminated – before its completion. We do seem to have clear judgments about such cases, and an adequate theory of causation needs to account for them. Philosophers argue about whether there is a way of doing so using counterfactuals.<sup>4</sup> My interest is reasons and rules, not causation, so I am going to avoid that debate. It is difficult enough to think of good examples of second-order reasons. Add that to the difficulty of thinking of examples of preemptive causation with reasons, and the cases left will be unusual. My main interest is the potential range of reasons with a counterfactual content.

### *Reasons to Act 'As If'*

I want to place second-order reasons in a broader context by relating them to other reasons that also have a counterfactual content. We can imagine reasons that are like second-order

---

<sup>3</sup> Tony Honoré, *Responsibility and Fault* (Hart 1999) 118.

<sup>4</sup> See, for example: Lewis 'Causation' (n 1); David Lewis, 'Postscripts to "Causation"' in *Philosophical Papers: Volume II* (OUP 1986); Jonathan Schaffer, 'Trumping Preemption' (2000) 97 *Journal of Philosophy* 165; John Collins, 'Preemptive Prevention' (2000) 97 *Journal of Philosophy* 223; David Coady, 'Preempting Preemption' in John Collins, Ned Hall, and LA Paul (eds), *Causation and Counterfactuals* (MIT Press 2004).

reasons except that what is counter to fact is not your believing there is a certain reason to  $\phi$ . What is counter to fact might be an event or a different type of omission (the non-occurrence of an event). For example, there might be a reason to make it the case that you both  $\phi$  and would not have  $\phi$ ed had it not been raining (when actually it is raining), or had it been raining (when actually it is not raining). We can also imagine reasons that are like second-order reasons except that they ask you to make it the case that *every* action of yours, and not just *some* action, is one you would not, or would have, performed had you not believed there is a certain reason. You satisfy such reasons only if none or all of your actions are ones you would have performed anyway.

These possibilities give us three dimensions of counterfactual reasons: they can be *positive* or *negative* in character; what is counter to fact can be an *omission* or an *event*; and they can apply to *some* or to *every* action. These dimensions are far from exhaustive, of course, and I intend this only as a partial typology of reasons with a counterfactual content. With three dimensions, there are eight types of reasons, shown in the chart below.

**Figure 1 – Types of Counterfactual Reasons**

	<i>Counterfactual</i>		<i>Application</i>		<i>Character</i>	
	<i>Omission</i>	<i>Event</i>	<i>Some action</i>	<i>Every action</i>	<i>Positive</i>	<i>Negative</i>
1	✓		✓		✓	
2	✓		✓			✓
3	✓			✓	✓	
4	✓			✓		✓
5		✓		✓		✓
6		✓		✓	✓	
7		✓	✓			✓
8		✓	✓		✓	

Positive second-order reasons are a sub-type of 1, distinguished primarily by the sort of omission involved. Likewise, negative second-order reasons are a sub-type of 2. What is interesting, and what the chart shows, is that second-order reasons are just a step or two away from a number of other reasons that also have a counterfactual content.

My interest is reasons of types 4 and 5. Since I am including forbearances as actions, it is possible to formulate the reasons more concisely than in the last paragraph. We can say, reasons of type 4 are reasons, for some actual event  $e$  and every action  $\phi$ , to  $\phi$  if you would have done so were  $e$  not to have occurred. Reasons of type 5 are reasons, for some non-actual event  $e$  and every action  $\phi$ , to  $\phi$  if you would have done so were  $e$  to have occurred. These reasons are still a bit of a mouthful, but all they really say is, do everything you would have done had this certain event not happened, or happened. Or, more concisely still, *act as if this*

certain event had not happened, or happened. I will call these *negative* and *positive reasons to act 'as if'*.

To be clear, negative and positive reasons to act 'as if' are *not* negative and positive second-order reasons. A second-order reason belongs under type 1 or 2, whereas a reason to act 'as if' belongs under type 4 or 5. Reasons to act 'as if' are more widely applicable, and range over more types of events, than second-order reasons.

### *Examples of Reasons to Act 'As If'*

There are lots of everyday examples of reasons to act 'as if'. You might have a reason to act as if you like your in-laws, even though you do not like them. Participants in fire drills, army training, and flight-test simulations have a reason to act as if they are in danger, to prepare them for when they actually are in danger.<sup>5</sup> Actors have a reason to act as if they feel what their characters feel, even when they do not, and lion tamers have a reason to act as if they were not afraid of lions, even when they are. Having overheard a snide remark about yourself at a party, it might be prudent to act as if you had not. To keep your children calm, you might have a reason to act as if you were not lost, even though you are.

Literature and film provide more exotic examples. In Stendhal's novel *The Red and the Black*, the main character, Julian, is deeply in love with Mathilde. After exchanging letters, they meet, and then again, but eventually Mathilde, sure of Julian's love, loses interest in him. For a long time he is crushed. Fortunately, Julian runs into a friend, Prince Korasoff, who advises him he can regain Mathilde's love indirectly, by pursuing someone else within her

---

<sup>5</sup> It would probably be more common to say the participants have a reason to act as if they *are* in danger, rather than a reason to act as if they *believe* they are in danger. But I consider the latter formulation to be more accurate, because it may be that if the participants were in danger they would not know (or believe) that.

social circle. Julian follows Korasoff's advice, and while at first Mathilde is puzzled by Julian's seeming indifference, she quickly becomes jealous of his attention towards her acquaintances, and finally she falls in love with him again. Julian, too, is in love, but to preserve her love, he must keep his hidden.

‘Ah,’ he said to himself, hearing the empty words his mouth had pronounced, as if he had been emitting strange noises, ‘if I could cover those pale cheeks with kisses, and you never felt them!’<sup>6</sup>

It is plausible to think that Julian has a reason to act as if he does not love Mathilde, even though he does. The reason is that, perversely, Mathilde will love Julian only if she believes that he does not love her, and acting as if he does not love her will make her believe that he actually does not.

The film *Life is Beautiful* is a similar example, except with respect to belief rather than love. Set in a World War Two concentration camp, the film shows a father who tries to hide the reality of the camp from his son by acting as if it is a game. In the game, the son can earn points by hiding from guards, not complaining, not asking for his mother, etc. The son comes to actually believe the camp is a game, and so is shielded from some of its psychological harms. He is also motivated to perform actions that will help keep him safe.

In *The Red and the Black*, Julian acts ‘as if’ to prompt a feeling in Mathilde. The father in *Life is Beautiful* acts ‘as if’ to prompt a belief in his son. In other situations, one person tries to prompt a feeling or a belief in another person by getting the *other* person to act ‘as if’. In his autobiography, Ben Franklin tells how he relied on this strategy to overcome his neighbour’s dislike of him:

I did not ... aim at gaining his favour by paying any servile respect to him but, after some time, took this other method. Having heard that he had in his

---

<sup>6</sup> Stendhal (Marie-Henri Beyle), *The Red and the Black* (first published 1830, translated by Burton Raffel, Modern Library 2003) 403. Jon Elster discusses Julian’s situation in *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge University Press 1983) 43ff.

library a certain very scarce and curious book, I wrote a note to him, expressing my desire of perusing that book, and requesting he would do me the favour of lending it to me for a few days. He sent it immediately, and I return'd it in about a week with another note, expressing strongly my sense of the favour. When we next met in the House, he spoke to me (which he had never done before), and with great civility; and he ever after manifested a readiness to serve me on all occasions, so that we became great friends, and our friendship continued to his death.<sup>7</sup>

Franklin did not try to gain his neighbour's affection directly. Instead, he cleverly brought it about that his neighbour acted as if he already liked Franklin. The tension between what the neighbour did, and how he felt, led to a change in his attitude toward Franklin.<sup>8</sup>

Ben Franklin engineered a change in his neighbour's actions as a way of changing his *neighbour's* attitudes. Sometimes, people have a reason to use the same strategy to change their *own* attitudes. For example, someone in a struggling marriage, who no longer loves his partner, might be wise to act as if he loves her, at least temporarily, with the idea that in time his feelings may come to reflect his actions. Hamlet gives similar advice when he tells his mother, Gertrude, how she can resist going to bed with his uncle:

Assume a virtue, if you have it not.  
That monster, custom, who all sense doth eat,  
Of habits devil, is angel yet in this,  
That to the use of actions fair and good  
He likewise gives a frock or livery,  
That aptly is put on. Refrain to-night,  
And that shall lend a kind of easiness  
To the next abstinence: the next more easy;  
For use almost can change the stamp of nature,  
And either master the devil, or throw him out  
With wondrous potency. (Act 3, Scene 4)

---

<sup>7</sup> Benjamin Franklin, *The Autobiography of Benjamin Franklin* (J Bigelow ed, Putnam's 1916) 216-217.

<sup>8</sup> The phenomenon involves what is known in the psychological literature as 'cognitive dissonance', that is, an inconsistency between your actions and beliefs which cannot be rationalized. There is empirical research which suggests that, sometimes, cognitive dissonance is resolved by a change in belief. See Leon Festinger and James Carlsmith, 'Cognitive Consequences of Forced Compliance' (1959) 58 *The Journal of Abnormal and Social Psychology* 203 for an early example of the psychological research, and Brian Huss, 'Three Challenges (and Three Replies) to the Ethics of Belief' (2009) 168 *Synthese* 249, 262ff for a discussion of its philosophical relevance.

Gertrude is not now virtuous; she finds it difficult, at least, to refrain from going to Hamlet's uncle. However, by acting as if she already were virtuous, Hamlet seems to say, she can change 'the stamp of nature' and thereby become virtuous.

It is sometimes claimed that belief can be induced in the same way. In *The Principles of Psychology*, William James writes:

[Gradually] our will can lead us to [a belief] by a very simple method: we need only in cold blood act as if the thing in question were real, and keep acting as if it were real, and it will infallibly end by growing into such a connection with our life that it will become real.<sup>9</sup>

Belief cannot be acquired at will, James says, but a belief can sometimes be acquired by acting as if you already had it. The 'thing in question', he writes, 'will become so knit with habit and emotion that our interests in it will be those which characterize belief'.<sup>10</sup>

Religious beliefs are among those which are said to be capable of bringing about in the way James describes. Apparently Talmud scholars 'preached not to require that people believe before they pray, only to do what is needed to get them to begin to pray; then they will come to believe in what and to whom they are praying'.<sup>11</sup> Pascal's Wager neatly combines the same idea with a practical motivation. Pascal thought it was more prudent to believe that God exists than to not believe it, but he recognized the difficulty of acquiring this belief directly. You cannot decide to believe God exists and, just like that, come to believe God exists. His suggestion is to instead act as if you already believed God exists:

You want to cure yourself of unbelief and you ask for the remedies? Learn from those who have been bound like you, and who now wager all they have .... Follow the way by which they began: by behaving just as if they

---

<sup>9</sup> William James, *The Principles of Psychology* (first published 1890, Cosimos 2007) 321.

<sup>10</sup> James, *The Principles of Psychology* (n 12) 321.

<sup>11</sup> Philip Zimbardo, *The Lucifer Effect* (Random House 2008) 449.

believed, taking holy water, having masses said, etc. That will make you believe quite naturally and according to your animal reactions.<sup>12</sup>

Rather than trying to directly acquire a belief in God, Pascal advises that you let your actions lead your beliefs.

### *Instrumental Reasons to Act 'As If'*

In all of the examples in the last sub-section, the reason to act 'as if' is an instrumental reason. The lion-tamer has a reason to act as if he is not afraid of lions because by doing so he will impress the audience or, less likely, intimidate the lion; the father in *Life is Beautiful* has a reason to act as if the camp is a game because by doing he will protect his son; Julian in *The Red and the Black* has a reason to act as if he does not love Mathilde because then she will come to love him; and so on. Perhaps there are non-instrumental reasons to act 'as if', but insofar as these are typical examples, they suggest that acting 'as if' is usually justified on instrumental grounds, when it is justified at all.

My focus later will be on acting as if you believe something you do not actually believe, so it is worth mentioning two instrumental justifications for doing so. First, acting as if you believe something you do not is sometimes a means of reducing the number of errors you commit, or the costs of your errors. For example, it is difficult to reliably distinguish dangerous pitbulls from harmless ones. Since it is much worse to be bitten than to avoid a harmless pitbull, it might be prudent to act as if you believe all pitbulls are dangerous. By doing so, you will increase the likelihood that, when you err, it will be by avoiding a harmless pitbull than by approaching a dangerous one.

---

<sup>12</sup> Blaise Pascal, *Pensées and Other Writings* (Honor Levi translator, OUP 1995) 155-6. See, for example: William James, *The Varieties of Religious Experience* (first published 1902, Harvard University Press 1985) 52: 'We can act as if there were a God; feel as if we were free; consider Nature as if she were full of special designs; lay plans as if we were immortal; and we find then that these words do make a genuine difference in our moral life.'

In other cases, acting as if you believe something you do not may be a means of saving time or labour. For example, you might have to decide, again and again, which bottle of wine to buy for dinner. You could perhaps figure out, on each occasion, which is the best bottle to buy, but doing so would require a great deal of research and thinking, both of which take time. On the other hand, if you act as if you believe a red wine from Italy is always best, even though you actually do not, you will consider fewer options and so decide more quickly, thereby saving time and effort.

Error-avoidance and efficiency in decision-making are, of course, two of the traditional instrumental reasons for having rules. In Chapter 10 I show this is more than just coincidence. Having a rule, insofar as it is instrumentally justified, is a means of acting as if you believe a normative proposition you do not.

### *Contexts and Reasons to Act 'As If'*

Most of the reasons to act 'as if' I have mentioned are reasons to act 'as if' in some contexts but not others. An actor does not have a reason to act as if he feels what his character does in all contexts; he has a reason to act that way on stage and in rehearsal. Julian in *The Red and the Black* has a reason to act as if he does not love Mathilde when he interacts with her, but not otherwise. When asking Prince Korasoff for advice, for example, he should be honest about how he feels. There might be a reason to act as if you believe all pitbulls are dangerous when faced with a pitbull, but not when debating the merits of a law banning pitbulls.

Sometimes, the context will be defined by your engagement in a particular type of activity, such as deliberating, arguing, or criticizing. For example, you might have a reason to act as if you believe red wine from Italy is always best, but only in the context of deliberating about which wine to buy. In this type of situation, we can make the context-relativity of the

reason clear by saying that you have a reason to deliberate about which wine to buy as if you believe a red wine from Italy is always best.

Another way of defining the context is by your *not* believing something. You might have a reason to act as if you believe pitbulls are always dangerous, except in contexts in which you believe a pitbull is not dangerous, or believe to a certain degree of confidence that it is not dangerous. Coming across a toothless, three-legged pitbull, you might be sure it is not dangerous, in which case the reason does not ask you to act as if you believe otherwise.

### *Causation and Counterfactuals, Again*

At the start of this section, I tried to use counterfactuals to explain what second-order reasons are reasons for. It mattered there whether there is an adequate counterfactual theory of causation, or at least one good enough for my purposes. If there is not, it would mean my account of the content of second-order reasons is not successful either. Since I began talking about reasons to act ‘as if’, I have left causation to the side and focused on counterfactuals. That has the advantage of limiting my reliance on counterfactual theories of causation. If there is no adequate counterfactual theory of causation, then reasons to act ‘as if’ are not analogous to second-order reasons in the way I have claimed they are. But there could still be reasons to act ‘as if’ and they would still have the features I have mentioned. That is all I suppose in the rest of this thesis.

Having drawn that line, it is worth briefly mentioning what reasons to act ‘as if’ are reasons for in terms of causation. A reason to act as if an event occurs, though it does not, is a reason to make it the case that nothing you do is caused by the non-occurrence of that event. A reason to act as if an event does not occur, though it does, is a reason to make it the case that nothing you do is caused by the occurrence of the event.

### *From Reasons to Acceptance*

In this section, I recast the content of second-order reasons in counterfactual terms. I then discussed in detail an analogous type of reason I called a reason to act ‘as if’. Reasons to act ‘as if’ are reasons to act in all respects as if some event occurred (or did not occur) even though actually it did not occur (or did occur). Typically, reasons to act ‘as if’ are instrumental reasons, and they favour acting ‘as if’ in some contexts but not others.

Going ahead, my interest is reasons to act as if you believe a proposition you do not actually believe. My immediate concern is with the means of satisfying such reasons. How can you bring yourself to act as if you believe  $p$  when actually you do not, and to do so in some contexts but not others? I try to answer this question in the next section. I discuss in some detail the distinction between belief and acceptance as it is drawn in the philosophy of action. After making a number of preliminary points, I show that accepting  $p$  brings you to act as if you believe  $p$ , and thus that accepting  $p$  is a means of satisfying a reason to act as if you believe  $p$ .

The motivation for the attention to reasons to act as if you believe what you do not, and to acceptance as an attitude that brings you into conformity with them, is that together these notions provide a way of making sense of rules. To anticipate the argument in later chapters, having a rule can be thought of as accepting a certain type of normative proposition. When having a rule is instrumentally justified, it is usually because it is a means of satisfying a reason to act as if you believe that proposition.

### 7.3 ACCEPTANCE AND BELIEF

Recently, a number of philosophers have contrasted belief with another propositional attitude, usually termed ‘acceptance’.<sup>13</sup> The details of their accounts vary, but there is a core of agreement, and that is enough for my purposes. In Robert Stalnaker’s formulation, ‘[t]o accept a proposition is to treat it as true for some reason. One ignores, at least temporarily, and perhaps in a limited context, the possibility that it is false’.<sup>14</sup> Above all it is the notion of *treating as true* that is central to acceptance. The usual reason to treat a proposition as true is that you believe it *is* true. As a result, what you believe you normally accept. However, you can also accept a proposition you do not believe. That would bring you to treat the proposition as true, and to act as if you believed it, though actually you do not.<sup>15</sup> It is also possible to not accept what you do believe.

In what follows, I state the main differences between belief and acceptance and then discuss some of the ways the attitudes are related. After taking a detour to specify some types of acceptances, I return to reasons to act ‘as if’. I conclude the section with a brief discussion of acceptance-based statements.

---

<sup>13</sup> The main works I draw on are by Robert Stalnaker: *Inquiry* (MIT Press 1987) 79-81; ‘Common Ground’ (2002) 25 *Linguistics and Philosophy* 701, 715-716; Jonathan Cohen: ‘Belief and Acceptance’, (1989) 98 *Mind* 367; *An Essay on Belief and Acceptance* (OUP 1995); and Michael Bratman, ‘Practical Reasoning and Acceptance in a Context’ (1992) 101 *Mind* 1. See also: Bas van Fraassen, *The Scientific Image* (OUP 1980); John Perry, ‘Belief and Acceptance’ (1980) 5 *Midwest Studies in Philosophy* 533; Edna Ullmann-Margalit and Avishai Margalit, ‘Holding True and Holding as True’ (1992) 92 *Synthese* 167; David Velleman, ‘On the Aim of Belief in *The Possibility of Practical Reason* (OUP 2000) 250; Keith Frankish, *Mind and Supermind* (Cambridge University Press 2004); Andrei Buckareff, ‘Acceptance and Deciding to Believe’, (2004) 29 *Journal of Philosophical Research* 173; Robert Audi, ‘Belief, Faith, and Acceptance’ (2008) 63 *International Journal of Philosophy of Religion* 87; Mark Sainsbury, ‘Fiction, and Acceptance-Relative Truth, Belief, and Assertion’ in Franck Lihoreau (ed), *Truth in Fiction* (Transaction Books 2011). The introduction and essays in Pascal Engel (ed), *Believing and Accepting* (Springer 2000) are also helpful.

<sup>14</sup> Stalnaker, ‘Common Ground’ (n 13) 716.

<sup>15</sup> See, for example: Stalnaker, *Inquiry* (n 13) 79-80: ‘To accept a proposition is to treat it as a true proposition in one way or another – to ignore, for the moment at least, the possibility that it is false. ... To accept a proposition is to act, in certain respects, as if one believed it.’

## *Distinguishing Belief and Acceptance*

This sub-section sets out three dimensions along which acceptance and belief differ. There may be others, but there is a consensus on these three.

*Context-independence versus context-dependence.* Belief, insofar as it is reasonable, is independent of context. According to Michael Bratman:

[A]t any one time a reasonable agent normally either believes something (to degree  $n$ ) or does not believe it (to that degree). She does not at the same time believe that  $p$  relative to one context but not relative to another.<sup>16</sup>

More precisely non-indexical belief is context-independent. You can believe ‘the book is here’ in one context but not another, since the context affects the content of the belief. Also, it is possible to have different attitudes towards a proposition depending on how it is presented.<sup>17</sup> Putting these complications aside, a reasonable person does not believe something relative to one context but not another. Of course, a person’s beliefs *can* change from one context to another. ‘But’, Bratman says, ‘such changes will likely get this person in trouble in his planning, and are not a case of *reasonably* believing that  $p$  in one context but not another’.<sup>18</sup>

Acceptance, unlike belief, can reasonably vary depending on the context. As Stalnaker describes it,

what a person accepts can be compartmentalized in a way in which what he believes cannot be. A person may accept something in one context, while rejecting it or suspending judgment in another. There need be no conflict that

---

<sup>16</sup> Bratman, ‘Practical Reasoning’ (n 13) 3. See also: Cohen, *Belief and Acceptance* (n 13) 13; Stalnaker, *Inquiry* (n 13) 80-81.

<sup>17</sup> Frankish, *Mind and Supermind* (n 13) 129; Bratman (n 13) 3 n 4.

<sup>18</sup> Bratman, ‘Practical Reasoning’ (n 13) 3. Emphasis in original. See also Stalnaker, *Inquiry* (n 13) 81: ‘[S]omething is wrong if I have separate incompatible beliefs for different circumstances. I cannot reasonably believe what I disbelieved yesterday without thinking that yesterday’s belief was mistaken’.

must be resolved when the difference is noticed, and he need not change his mind when he moves from one context to the other.<sup>19</sup>

It would be unreasonable to have different sets of beliefs for different circumstances; not so with respect to acceptance. You can accept something today, and not accept it tomorrow, without regarding your earlier acceptance as mistaken.

The terms ‘context-dependent’ and ‘context-relative’ are often awkward to use, so I will introduce a new term. I will say that the contexts (or types of contexts) relative to which an acceptance is held are the contexts (or types of contexts) in which the acceptance is *operative*. So, if you accept  $p$  relative to context  $C$  but not relative to context  $C'$ , then your acceptance is operative in  $C$  but inoperative in  $C'$ .

The following example illustrates the limited operation of acceptance. Sales assistants are required to accept that ‘the customer is always right’. No one actually believes this, of course, including sales assistants. Some customers are right, but others make honest mistakes, and others lie and steal. Nonetheless, sales assistants must treat it as true that the customer is always right. As Mark Sainsbury says:

[T]o do their job properly, in certain circumstances, sales assistants have to act as if they believed [that the customer is always right]. That means: listening carefully to what the customer says, not challenging it, making remedial proposals that take for granted the customer’s story, and so on.<sup>20</sup>

The sales assistants’ acceptance will be operative during store hours and when interacting with customers. During those hours, they will act as if they believe the customer is always right. But ‘[a]fter hours, it’s quite alright for sales assistants to tell one another how wrong the customers are’.<sup>21</sup>

---

<sup>19</sup> Stalnaker, *Inquiry* (n 13) 80-81

<sup>20</sup> Sainsbury, ‘Fiction and Acceptance-Relative Truth’ (n 13) 149. For a similar discussion of the same example, see Ullmann-Margalit and Margalit, ‘Holding True and Holding as True’ (n 13) 170-171.

<sup>21</sup> Sainsbury, ‘Fiction and Acceptance-Relative Truth’ (n 13) 149.

*Truth versus Success.* Belief that is reasonable is normally shaped by evidence and a concern for the truth of what is believed. Hence the metaphor: belief aims at truth. Why suppose that belief aims at truth in a way that other attitudes do not? Bernard Williams offers several reasons.<sup>22</sup> First, it is appropriate to assess the truth or falsity of the content of belief, but inappropriate to do so with respect to most other attitudes. If you believe it is raining, it is appropriate to ask whether the belief is true or false; not so with respect to an intention to take your umbrella. Second, to believe *p* is to believe that *p* is true. If you believe it is raining, that is the same thing as believing that it is true it is raining. Third, if you say ‘I believe *p*’, you claim that *p* is true. It means you are committed to the truth of *p*. Of course some beliefs are shaped not by a concern for the truth but by what you wish were true. In that case, though, you are open to criticism for wishful thinking, self-deceiving, or something similar.

Acceptance, in contrast, does not aim at truth. Rather, it aims at success, measured in practical terms.<sup>23</sup> It is shaped mainly by practical considerations that are not themselves evidence for the truth of what is accepted. For that reason, you can reasonably accept in a context what you do not believe, and even what you believe is false. In the example above, the sales assistants do not believe the customer is always right; indeed, they almost certainly believe it is false. Nonetheless, they accept the customer is always right, when interacting with customers, and do so because it is a requirement of their job.

Consider another example, this time of a therapist who accepts what she believes is false in order to better treat a patient. The example is Sainsbury’s.

---

<sup>22</sup> Bernard Williams, ‘Deciding to Believe’ in *Problems of the Self* (Cambridge University Press 1976) 137; see also Bratman, ‘Practical Reasoning’ (n 13) 3.

<sup>23</sup> Bratman, ‘Practical Reasoning’ (n 13) 9; Cohen, *Belief and Acceptance* (n 13) 25, 100-103; Stalnaker, *Inquiry* (n 13) 81.

A disturbed patient is recounting his (entirely fictitious) early history to his therapist:

When I was young, I played the violin. I performed Beethoven's sonata in E flat at the Wigmore Hall.

The therapist knows this is false, but decides it's best to roll with her patient's delusions and says:

Did you play an encore?<sup>24</sup>

In asking the question, 'Did you play an encore?', the therapist is treating it as true that her patient was a concert violinist.<sup>25</sup> Why does she do so? Not because she believes it is true. It is plainly false. Rather, she treats it as true to better bring out the patient's fantasies, as a means of providing him with more effective treatment. She accepts that her patient is a concert violinist for practical reasons, not truth-related ones.

Stalnaker gives a similar example of a scientist who accepts certain propositions to simplify his inquiries.

Particles or planets may be treated as mass points, the atmosphere may be assumed to be a vacuum, consumers or governments may be thought of as rational. Of course in other inquiries these same assumptions might greatly distort the results, but the scientist might be in a position to know that in his inquiry they would not. The scientist does not, of course, *believe* the propositions he accepts, but he acts, in a limited context, as if he believed them in order to further his inquiry.<sup>26</sup>

The scientist accepts what he does not believe (indeed, what he believes to be false) because it will help him with his inquiries. In this respect, he is guided by his goals, not by the evidence. In another context – say, when the composition of the atmosphere is under consideration – his goals will be different, and what he treats as true will change accordingly.

---

<sup>24</sup> Sainsbury, 'Fiction and Acceptance-Relative Truth' (n 13) 153-154.

<sup>25</sup> Sainsbury, 'Fiction and Acceptance-Relative Truth' (n 13) 154: '[I]t's hard to see how one could explain the therapist's mental states without introducing this content. That's because it's presupposed by the question, and the therapist must accordingly be at least implicitly aware that this is so.'

<sup>26</sup> Stalnaker, *Inquiry* (n 13) 93.

*Active versus Passive.* Belief is not normally subject to direct, voluntary control. You cannot decide to believe *p* and, just like that, come to believe *p*. Rather, the world as it is, as well as demands of reasonableness, shape your beliefs. Jonathan Cohen says:

Beliefs are said to come over you, arise in you, or grow on you, like anger or affection does. You cannot don, raise, or grow them yourself. You can plant them in others, but not in yourself. You may decide to believe a friend, i.e. to trust his word. You may even decide to believe *in* him, i.e. to have confidence in his abilities. You may decide also whether to let your mind dwell on a belief or to try to keep your mind off it, whether to check the evidence for it or to disregard it. But you cannot decide to believe *that* it will rain tomorrow or *that* it will not.<sup>27</sup>

The passiveness of belief is not a contingent feature of it, Williams argued.<sup>28</sup> It flows from the way that belief aims at truth. If you could simply choose what to believe at will, you could do so irrespective of the truth of the content of the belief. But if you knew that one of your attitudes was not shaped by a concern with the truth of its content, you could not regard it as a belief.

It is important to keep in mind that what people lack is *direct* control over their beliefs. Pascal, for example, thought that engaging in the practices associated with the belief that God exists could bring you to acquire that belief. If you did as Pascal suggested – took holy water, had masses said, etc. – you might come to believe God exists, and it would have been the result of a voluntary action. However, the belief would not have been *directly* acquired. Instead it was acquired indirectly, by means of taking holy water, saying masses, etc.<sup>29</sup> Indeed, the difficulty of acquiring the belief directly is why Pascal proposes such an elaborate strategy.

Take another example. Suppose your friends are arguing about where to go for dinner. To resolve the disagreement, they ask you what you think. You must make up your mind,

---

<sup>27</sup> Cohen, *Belief and Acceptance* (n 13) 21.

<sup>28</sup> Williams, 'Deciding to Believe' (n 22); see also Bratman, 'Practical Reasoning' (n 13) 3. The argument has its critics. However, the conclusion – that direct doxastic voluntarism is false – has wide support, and that is what matters for my purposes.

<sup>29</sup> Cohen, *Belief and Acceptance* (n 13) 21.

even if it was divided before. That will not be something that just happens. It will be something you actively do. Nor, to be clear, is making up your mind just initiating a process of reasoning in the hope that it will lead to a belief. It means putting a relatively definitive end to your reasoning.<sup>30</sup> It appears, then, that by making up your mind you perform an action that produces a belief. In that sense, belief at least sometimes involves a direct, voluntary element. However, while you can voluntarily make up your mind, you cannot voluntarily make up your mind in a particular way. When your friends ask you to resolve their disagreement, you force yourself to take a view, but not to take the view of a certain friend. You do not directly and voluntarily come to have a belief with a certain content.

Acceptance, in contrast, is subject to direct, voluntary control, in the relevant sense. You can choose what to accept in a context and bring it about that you have that attitude, and you do not need to use a means to do so as you would to acquire a belief. Sainsbury's therapist, say, decides to treat it as true that her patient was a concert violinist, and then does so, without taking any step in between. Likewise, the sales assistants are able to accept that the customer is always right by deciding to do so. They do not need to hypnotize themselves, seek out evidence that the customer is always right, or employ any of the other indirect methods that would be necessary to induce such a belief.

To summarize, acceptance and belief tend to go hand in hand, but the two are different attitudes. Whereas belief is context-independent, responsive to evidence, and out of your direct, voluntary control, acceptance is none of these things. It is context-dependent, responsive to practical reasons, and under your direct, voluntary control.

---

<sup>30</sup> Frankish, *Mind and Supermind* (n 13) 21-22.

## *Belief and Acceptance Related*

I have said how belief and acceptance are different; now I want to say how they are related. It is uncontroversial in the literature on belief and acceptance that you can accept  $p$  but not believe  $p$ . It is also uncontroversial that you can believe  $p$  but not accept  $p$  in a particular context. That is, acceptance does not entail belief, nor does belief entail acceptance in a particular context. (Whether belief entails acceptance in *some* context is unclear. It is also unimportant for my purposes.)

Acceptance and belief can diverge, but normally they do not. Normally, what you accept in a given context is what you believe. That is because accepting a proposition is treating it as true and, as Stalnaker says, ‘the simplest reason to treat a proposition as true is that one believes it *is* true’.<sup>31</sup> To use Bratman’s term, by *default* you accept in a context whatever you believe.<sup>32</sup> For example, most of what the sales assistants accept is due to what they believe. Their acceptance that the customer is always right marks the exception, not the rule.

Some further terminology would be helpful at this point. Suppose you accept that  $p$ . Your acceptance might depend on your continued belief that  $p$ , or not. Call your acceptance *belief-dependent* in the first case and *belief-independent* in the second case. It is possible to both believe  $p$  and to independently accept  $p$ . What is the significance of the independence of the acceptance? For one thing, it provides added certainty that you will treat  $p$  as true. Even if you were confronted with persuasive evidence that not- $p$ , which would lead you to abandon your belief that  $p$ , you would not for that reason cease to accept  $p$ , or cease to treat  $p$  as true.<sup>33</sup>

---

<sup>31</sup> Stalnaker, ‘Common Ground’ (n 13) 716.

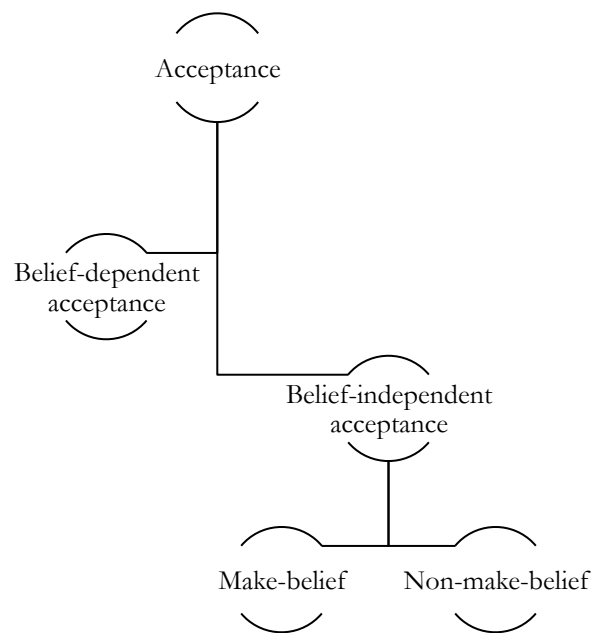
<sup>32</sup> Bratman, ‘Practical Reasoning’ (n 13) 10

<sup>33</sup> Philip Pettit makes a similar point about precautionary assumptions: ‘Hope and its Place in the Mind’ (2004) 592 *Annals of the American Academy of Political and Social Science* 152. A precautionary assumption still has a role to play, even if it changes

Your acceptance could be belief-independent either because you do not believe  $p$ , or for some other reason. Call the acceptance *make-belief* in the first case and *non-make-belief* in the second case. (I mean ‘make-belief’ in the sense of the simulation of a belief not actually possessed, not in the technical sense it sometimes takes in philosophy.<sup>34</sup>) All the examples I have given of acceptance so far are of make-belief. The sales assistants make-believe that the customer is always right; the therapist make-believes that her patient is a concert violinist; and the scientist make-believes that the atmosphere is a vacuum. They all accept what they do not believe. The figure to the side summarizes the relationships between the types of acceptances.

The conceptual difference between make-belief and belief is clear, even though the first tends to bleed into the second. Indeed, a common reason to make-believe  $p$  is to bring yourself to believe  $p$ .<sup>35</sup> Earlier

**Figure 2 – Types of Acceptances**




---

your beliefs, because it guards against the danger ‘of going along initially with that cautious belief but then fluctuating in your resolve as the belief waxes and wanes with incoming evidence. Tying yourself to the precautionary assumption will keep you on track ... in a way that trusting yourself to the revised belief might not’ (155).

<sup>34</sup> See for example: Kendall Walton’s *Mimesis and Make-Believe* (Harvard University Press 1990) and ‘Metaphor and Prop Oriented Make-Believe’ (1993) 1 *European Journal of Philosophy* 39; Velleman, ‘On the Aim of Belief’ (n 13) 256-263.

<sup>35</sup> See Tamar Gendler, ‘On the Relation Between Pretense and Belief’ in Domenic Lopes and Matthew Kieran (eds), *Imagination, Philosophy and the Arts* (Routledge 2003) 131-136 for an overview of how pretending that  $p$  sometimes leads to the belief that  $p$ , with special emphasis on children’s make-believe.

I mentioned how Pascal thought that you could acquire a belief in God by engaging in the practices associated with the belief. Make-believing that God exists brings you to do just that.

As Cohen says:

[A]cceptance that  $p$  very often *causes* belief that  $p$  – in the long run if not in the short run. Thus Pascal urged his readers to act in accordance with Catholic doctrine – that is, to take it as a premise for their practical decisions – since such a policy would tend eventually to induce belief that the doctrine was true.<sup>36</sup>

By making Pascal's wager, you gamble that your make-belief will lead to actual belief. Suppose it does. You now believe that God exists, and you believe that because you accepted it. Moreover, we can imagine you later ceasing to believe that God exists but continuing to accept that God exists, perhaps in the hope of reacquiring the belief at some point. Your acceptance will coincide with a belief with the same content, but it will still be what I termed 'belief-independent'.

### *Make-Belief and Reasons to Act 'As If'*

In the last section, I discussed reasons to act 'as if'. Some reasons to act 'as if' are reasons to act as if you believed something you actually do not. I did not say how you can bring yourself to satisfy such a reason. Now I am able to.

Normally, if you believe  $p$ , you will accept  $p$  in  $C$ . That will lead you to treat  $p$  as true in  $C$ . What does it mean to treat  $p$  as true in  $C$ ? It depends on what ' $p$ ' and ' $C$ ' stand for. Sainsbury's sales assistants treat it as true that the customer is always right by interacting with customers in a certain way when they are on duty. Stalnaker's scientist treats it as true that the

---

<sup>36</sup> Cohen, *Belief and Acceptance* (n 13) 18. Buckareff, 'Acceptance and Deciding to Believe' (n 13) 177 discusses Pascal's Wager in similar terms: 'By wagering one accepts (that God exists) for practical reasons. Accepting that God exists guides the agent who desires to believe that God exists to perform actions that will eventually give rise to belief.'

atmosphere is a vacuum by using it as basis for inference and a guide to further inquiry. In general, to treat a proposition as true means acting in ways that are appropriate given the truth of the proposition and the context one is in.

If you do not believe  $p$ , but you make-believe  $p$  in  $C$ , you will treat  $p$  as true in  $C$ . You will act in  $C$  in the ways you normally would act were you to believe  $p$ . In other words, make-believing  $p$  in  $C$  leads you to act in  $C$  as if you believe  $p$ . This is why Sainsbury says the sales assistants who make-believe that the customer is always right ‘act as if they believed’ that. It is why Stalnaker says that the scientist who make-believes that the atmosphere is a vacuum ‘acts ... as if he believed’ it is a vacuum. And it is why make-believing  $p$  in  $C$  is normally a way to satisfy a reason to act in  $C$  as if you believe  $p$ . To return to an earlier example, make-believing that all pitbulls are dangerous is normally a way to satisfy a reason to act as if you believe they are all dangerous. Likewise, make-believing the concentration camp is just a game is normally a way to satisfy a reason to act as if you believe that.

I have been careful to say that *normally* make-believing something in the appropriate context is a means of satisfying a reason to act as if you believed it. The qualification is necessary because it is only normally the case that you accept what you believe. It is possible that you would not accept  $p$  were you to believe  $p$ , in which case make-believing  $p$  (as a form of accepting  $p$ ) would not lead you to act as if you believed  $p$ . More plainly, if you would not treat  $p$  as true if you believed it, then making it so you treat  $p$  as true is not a way to act as if you believed it.

How important are these ‘abnormal’ cases? Not very. They are cases in which you would not act any differently whether you believed  $p$  or not. It is difficult, at least, to see why you would have a reason to act as if you believed  $p$  if it would not make any difference to what you do. For that reason, I do not regard these cases as important. It is at least very

unlikely that you will have a reason to act in  $C$  as if you believed  $p$  but make-believing  $p$  in  $C$  is not a means of satisfying the reason.

One final point: I have discussed how make-believe is a means of satisfying a reason to act as if you believe something you actually do not. What about a reason to act as if you do *not* believe something you actually do believe? All that is needed is for you to *not* accept whatever it is you believe. That will lead you to *not* treat it as true, just as you normally would were you not to believe it. (There are abnormal cases where, were you not to believe  $p$ , you would make-believe  $p$ , but we can set such cases aside.)

### *Acceptance and Statements*

Ultimately I will argue that rule-possession is based on a kind of acceptance, and at that point it will be important whether acceptance can lead people to make the kind of communications associated with rule-possession. At this point I wish to show that make-believing  $p$  can lead you to make the same or similar communications you would make were you to believe  $p$ . Suppose the manager of a store tells his sales assistants: ‘The customer is always right’. By making the utterance, the manager performs a speech act. What kind of speech act? It might be an assertion. There is a problem, however. An assertion is normally thought to imply that the speaker believes what he says. The manager, on the other hand, accepts that the customer is always right, but he does not believe that. He is not even representing that he believes that, and it would be unreasonable for the sales assistants to think that he does. If an assertion really does imply a belief in what is said, then the manager did not make an assertion. More generally, verbal expressions of make-beliefs would not count as assertions.

There are then two possibilities. First, belief may not be as essential to assertion as most people think. Maybe they just tend to go together. In that case, the fact the manager

does not believe the customer is always right does not disqualify him from asserting that. Among philosophers who write on acceptance, Cohen is the most prominent advocate of this view. Taking ‘assertions in the ordinary, non-technical sense of the word, in which they involve the declaration of a claim and some expectation of its recognition or insistence on its validity,’ Cohen claims that

there is no reason at all why an assertion that  $p$  should normally imply that the speaker believes that  $p$ . He may well be insisting on recognition of his claim that  $p$  because he wants people to know that he accepts that  $p$ , though he lacks as yet – and perhaps will never have – any corresponding belief.<sup>37</sup>

Cohen sees no oddity in sentences like ‘it is raining but I do not believe it is raining’, because he does not think saying ‘it is raining’ implies one believes it is raining. It may simply express an acceptance, and there is nothing troubling or strange about someone accepting it is raining but not believing it is raining.<sup>38</sup>

The second option is to deny that a verbal expression of an acceptance is an assertion. It would be some other type of speech act, analogous to what an actor says on stage.<sup>39</sup> In this case, by saying ‘the customer is always right’, the manager would not be making an assertion, since he lacks the necessary belief. Instead, he would be performing another type of speech act. For my purposes, it does not matter much which option is better. The important point is just that the manager performs a speech act that at least outwardly resembles an assertion, and thus outwardly he acts as if he believes the customer is always right.

---

<sup>37</sup> Cohen, *Belief and Acceptance* (n 13) 71.

<sup>38</sup> Cohen, *Belief and Acceptance* (n 13) 72. For a discussion of acceptance and Moore’s Paradox, see Pascal Engel, ‘Believing, Holding True, and Accepting’ (1998) 1 *Philosophical Explorations* 140, 140-143, as well as ‘Moore’s Paradox’ by Laurence Goldstein and ‘On “Moore’s Paradox”’ by Robert Stalnaker, both in Engel (ed), *Believing and Accepting* (n 13).

<sup>39</sup> Nicolas Unwin ‘What Does it Mean to Aim at Truth?’ (2003) 40 *American Philosophical Quarterly* 91, 93: ‘[E]xpressions of acceptances ... like quotations and theatrical utterances ... should perhaps be re-classified as indicative speech acts of a different kind.’

## *Hartian Acceptance*

In Section 4.5 I said that Hart calls the attitude in which having a rule consists as ‘acceptance’ or ‘acceptance of a rule’. And now I have introduced an attitude with the same name. Are they the same attitude? There are certainly similarities. Hart’s acceptance is responsive to practical reasons; so is the kind of acceptance I have been talking about. Hart’s acceptance may be ‘voluntarily acquired’. If that means it is under your *direct* voluntary control, then the two attitudes share this feature, too. Hart does not discuss context-dependence, at least not directly, but he says nothing to suggest that the attitude he has in mind is *not* context-dependent. The same is true of the relationship acceptance bears to belief – Hart does not comment on it. So Hart’s notion of acceptance and the notion I have been talking about have at least some of the same features, and there are no explicit differences.

There is still some reason to think that these are different attitudes. For one thing, Hart seems to use the term acceptance as it is ordinarily used, whereas in the philosophy of action acceptance is an avowedly technical term, and the claims made about it are not meant to reflect ordinary usage.<sup>40</sup> More importantly, the kind of acceptance that Bratman, Stalnaker, and Cohen are talking about is a propositional attitude. Hart, in contrast, repeatedly refers to the acceptance of a rule or of a standard, and neither rules nor standards are propositions.<sup>41</sup> I return to these issues in Chapter 9, where I propose a way of reading Hart that would smooth over this seeming dissimilarity. For now, the point to keep in mind is that the way I understand acceptance is not necessarily incompatible with the way Hart understands it.

---

<sup>40</sup> See, e.g., Stalnaker, *Inquiry* (n 13) 79: ‘Acceptance is a technical term: claims I make about acceptance are not intended as part of an analysis of a term from common usage’. Many other philosophers who write on acceptance endorse the claims Stalnaker makes about acceptance.

<sup>41</sup> The content of a rule is a proposition, and so might be the content of a standard, but the rule or standard is not itself a proposition. It does not make sense, for instance, to say that a rule is true or false.

### *Summary*

To accept a proposition in a context is to be disposed to treat the proposition as true in that context. By default, you accept in a given context all and only what you believe. Beliefs do not vary from context to context. They are shaped by evidence and are beyond your direct, voluntary control. Ultimately, though, it is up to you what to accept in a given context and, if you are reasonable, your choices will be guided by practical reasons. When you accept what you do not believe, the acceptance is a make-belief. Make-believing something leads you to act as if you believed it, even though you do not, and is thus a means of satisfying a reason to act as if you believed it. Among other actions, make-believing something leads you to make the statements you would have made were you to have a belief with the same content as your make-belief.

## 7.4 CONCLUSION

The last several chapters were about rules. For a society to have a rule is for its members to have a certain attitude, which leads them to act as if they believe it ought to be complied with. In this chapter, I discussed the variety of reasons to act ‘as if’ and how a kind of acceptance may be the attitude we are looking for. In later chapters, I set out an acceptance-based theory of rule-possession and a make-belief-based understanding of why we ought to have rules. Since acceptance is the cornerstone of the chapters to come, I want to add to the evidence that it is what I have claimed. This is the main purpose of the next chapter, which illustrates the role of acceptance in the law of evidence.

# 8 Acceptance and Evidence Law

---

## 8.1 INTRODUCTION

Acceptance can seem too good to be true. Are people really able to accept whatever they want and nothing they do not? Can they so easily decide what to treat as true? For my purposes, the best response to such scepticism is that, in the eyes of the law, doing so is not just possible but often required. The clearest examples are from evidence law. Some rules of evidence law require judges and juries to presume a proposition whether or not they believe it, and sometimes even when they know it is false. Other rules require judges and juries not to deliberate based on what they do believe, most commonly when they are exposed to inadmissible evidence. In substance these are requirements to make it the case that one accepts, or does not accept, certain propositions in the context of one's deliberations at trial, regardless of what one actually believes.

This short chapter is meant as a bridge between the general discussion of acceptance in the last chapter and the discussion of acceptance and rules in the next chapter. It has two subsidiary aims. The first is to show that acceptance already has a place in law. That removes some of the mystery around acceptance and helps to show that the theory of rules I argue for is more familiar than it might seem. Both Section 3, which is about exclusionary rules, and Section 4, which is about presumptions, advance this first aim. Section 4 also advances a second aim, which is to lay the groundwork for an analogy between rules and presumptions, which I develop in the next chapter. The chapter is not meant as a contribution to evidence law theory, and the claims I make about evidence law are intended to be uncontroversial.

## 8.2 BASICS OF EVIDENCE LAW

In any trial, the plaintiff or prosecutor makes a claim or charge, and the defendant may make a defence. There are essential facts to any claim, charge, or defence, determined by substantive law. The parties usually disagree about some of these facts. One party will allege one thing and the other party will dispute it. The essential facts that are in dispute are known as the *facts in issue*. The *tribunal of fact* – the jury, or the judge if he is sitting alone – is in charge of deciding which of the facts in issue obtain. The law of evidence is roughly the law that governs how and on what basis the tribunal may, must, or must not reach that decision. The general rule is that the judge or jury, whichever it happens to be, must deliberate on the basis of the evidence and nothing else.<sup>1</sup> For that reason, it matters a great deal what the parties may introduce as evidence. This is for the *tribunal of law* – the judge, always – to decide. Information that may be introduced (or presented) to the tribunal of fact is *admissible* as evidence; information that may not be introduced is *inadmissible* as evidence. Only information that is relevant to a fact in issue – that is, that tends to show that the existence of such a fact is more or less likely – is admissible.<sup>2</sup> Information that is relevant is generally admissible, but not always. As Thayer says,

(1) ... without any exception, nothing which is not, or is not supposed to be, logically relevant is admissible; and (2) ... subject to many exceptions and qualifications, whatever is logically relevant is admissible ....<sup>3</sup>

---

<sup>1</sup> *Goold v Evans & Co* (1951) 2 TLR 1189 (CA) (Denning LJ): 'It is a fundamental principle of our law that a judge must act on the evidence before him and not on outside information'. See also the *Manual of Civil Model Jury Directions* issued by the U.S. Courts for the Ninth Circuit (2007) < <http://www.lb9.uscourts.gov/> > accessed 10 July 2011, §1.7: 'In reaching your verdict, you may consider only the testimony and exhibits received into evidence. ... Anything you may have seen or heard when the court was not in session is not evidence. You are to decide the case solely on the evidence received at trial'.

<sup>2</sup> See, for example, James Fitzjames Stephen, *A Digest of the Law of Evidence* (12<sup>th</sup> edn, Macmillan 1946) Article 1: relevant means 'any two facts to which it is applied are so related to each other that according to the common course of events one either taken by itself or in connection with other facts proves or renders probable the past, present, or future existence or non-existence of the other.'

<sup>3</sup> James Bradley Thayer, *A Preliminary Treatise of Evidence at the Common Law* (Little, Brown, & Co 1898) 266.

Much of the law of evidence consists of rules about whether information that is relevant is also admissible.

### 8.3 EXCLUSION

Judges and jurors are generally required to deliberate based on the evidence and only on the evidence. However, it is often a matter of dispute between the parties whether some information is admissible as evidence. One party seeks to have it admitted and the other resists. The judge, as tribunal of law, must rule on its admissibility. When the judge is sitting with a jury, the procedure is straightforward. Normally the jury leaves the courtroom and in their absence the judge hears argument from the parties about whether the information is admissible.<sup>4</sup> The hearing is a *voir dire*, a ‘trial within a trial’. To help him determine the information’s admissibility, the judge may consider the disputed information and receive evidence that goes to its admissibility. If he rules that the evidence is inadmissible, it is excluded, that is, the party offering the evidence is prevented from presenting it to the jury. The information is said to be *perceptually excluded*; it amounts to a *prohibition of proof*.<sup>5</sup>

Things are more complicated when the judge is sitting alone. He may be able to rule on the admissibility of the information without considering it, at least in detail. Often, though, that is impractical. It is hard, for instance, to decide whether an accused confessed – roughly, gave a statement against his interests – involuntarily, and hence inadmissibly, without considering the content of the statement.<sup>6</sup> In such cases, the judge will have to consider the

---

<sup>4</sup> This is the usual practice, but in *R v Reynolds* (1950) 1 KB 606 the Court of Criminal Appeal ruled that the jury should be present during a hearing to determine a child’s competency to testify. The jury would not hear anything potentially inadmissible, and would be left in a better position to assess the child’s truthfulness, if he or she were permitted to testify.

<sup>5</sup> Hock Lai Ho, *A Philosophy of Evidence Law* (OUP 2008) 43.

<sup>6</sup> Section 76 of the Police and Criminal Evidence Act 1984 defines a confession as ‘any statement wholly or partly adverse to the person who made it, whether made to a person in authority or not and whether made in words or otherwise.’

information to rule on its admissibility. If it turns out to be inadmissible, then the judge will ‘exclude’ it. However, it is too late to prevent it from being presented to the tribunal of fact – i.e., to the judge – as it could be if he were sitting with a jury. Instead, the judge is supposed to ‘exclude from his mind’ or ‘disregard’ the inadmissible information when, as tribunal of fact, he determines what happened.<sup>7</sup>

Something similar may be required of the jury when it is inadvertently exposed to inadmissible evidence: for example, when a witness reveals something he was not allowed to reveal. If the information is especially damaging, the judge may order a mistrial. Most of the time, though, the judge will just direct the jury to do what he would do were he sitting alone, i.e., disregard or mentally exclude what was heard or seen.<sup>8</sup> For example, here is a selection from the model jury instructions for the U.S. Sixth Circuit:

During the trial I ... ruled that you could not see some of the exhibits that the lawyers wanted you to see. And sometimes I ordered you to disregard things that you saw or heard, or I struck things from the record. You must completely ignore all of these things. Do not even think about them. Do not speculate about what a witness might have said or what an exhibit might have shown. These things are not evidence, and you are bound by your oath not to let them influence your decision in any way.<sup>9</sup>

Just as the jury is required not to deliberate based on anything they saw or heard outside of court, they are required to ignore some things they heard in court but should not have. The information is not perceptually excluded – it is too late for that – but rather *excluded from deliberation*.

---

<sup>7</sup> Raymond Emson, *Evidence* (5<sup>th</sup> edn, Palgrave Macmillan 2010) 21; Peter Murphy, *Murphy on Evidence* (11<sup>th</sup> edn, OUP 2009) 65.

<sup>8</sup> See, for example, *R v Lawson and others* (2007) 1 Cr App R 277 (CA). The defendants were charged with importing cannabis resin, and claimed in their defence they thought they were importing tobacco. The trial judge ruled that the jury should not be made aware an acquaintance of the defendants had been convicted for drug dealing. Later, the judge mistakenly revealed that fact to the jury. Instead of discharging the jury, he simply directed them to ‘ignore’ the evidence. The decision was upheld.

<sup>9</sup> *Pattern Criminal Jury Instructions* (2009) issued by the U.S. Courts for the U.S. Sixth Circuit <[http://www.ca6.uscourts.gov/internet/crim\\_jury\\_insts.htm](http://www.ca6.uscourts.gov/internet/crim_jury_insts.htm)> accessed 10 July 2011, §1.04.

Exclusion from deliberation is more fundamental than perceptual exclusion, the latter being a means of ensuring the former.

The reason why we exclude a certain matter from a fact-finder's perception or bar a party from adducing it as evidence is because we wish to prevent it from influencing deliberation. 'Perceptual exclusion' and 'exclusion as prohibition of proof' are merely means of enforcing 'exclusion from deliberation'. They are, undoubtedly, very effective means; indeed, they guarantee success: if the jury is kept in the dark about the information by preventing a party from presenting it, it is certain the information will not figure in their deliberations.<sup>10</sup>

When perceptual exclusion fails, jurors are relied on to achieve its aim by other means. They are required to simply make it the case that they do not deliberate based on certain information. To be clear, they are not required to abandon certain of their beliefs. A judge is not expected to stop believing that the accused confessed, for example. Rather, they are required to make it the case that something they *do* believe and which is relevant to a fact in issue does not figure in their deliberations – but not all their deliberations, just the ones at trial. Moreover, there is a clear assumption that this is under the direct control of the judge or jury; it is something they can decide to do and make happen.<sup>11</sup>

In all of these respects, what judges and jurors are required to do is to make it the case that they do not accept certain things they might believe in the context of their deliberations at trial. The judge who rules that the accused's confession is inadmissible, for example, must not accept that the accused confessed when he deliberates about whether the accused committed the crime. Since what one accepts in a context is under one's direct

---

<sup>10</sup> Ho, *Philosophy of Evidence Law* (n 5) 43.

<sup>11</sup> There are some notable exceptions. Justice Jackson of the US Supreme Court, for example, claimed that all lawyers know 'the naïve assumption that prejudicial effects can be overcome by instructions to the jury' to be 'unmitigated fiction': *Krulwich v United States* (1949) 336 US 440, 453. For overviews of the empirical research on the effectiveness of jury instructions, see: Michael Nietzel, Denis McCarthy, and Monica Kerr, 'Juries: The Current State of the Empirical Literature' in Ronald Roesch, Stephen Hart, and James Ogloff (eds), *Psychology and Law: The State of the Discipline* (Kluwer 1999) 33-37; Ehud Guttel, 'Overcorrection' (2004) 93 *Georgia Law Journal* 241; Robert MacCoun, 'Inside the Black Box: What Empirical Research Tells Us About Decision-Making by Civil Juries' in Robert Litan (ed), *Verdict: Assessing the Civil Jury System* (Brookings Institution 1993). In general, the empirical research does not distinguish the *ability* of the jury to exclude information from their deliberations from their *willingness* to do so, where only the former is relevant to my argument. For that reason, I do not regard studies that show jury instructions are sometimes ineffective as ground for objection.

control, the judge is able to make it the case that he does not accept that the accused confessed without using a means to do so. The judge will continue to believe that the accused confessed, of course, but he will not treat it as true that the accused confessed in his deliberations.

## 8.4 PRESUMPTIONS AND FICTIONS

### *Presumption and Acceptance*

As a rule, the tribunal of fact is required to reach a decision about the facts in issue based only on the evidence introduced by the parties. Everyday experience is an exception.<sup>12</sup> Another exception is that one party can formally admit a fact that the other party would otherwise be obligated to prove. The admission conclusively establishes that fact, without any need for evidence of it. My interest at the moment is a third exception: presumptions.

On occasion the law says that the judge or jury may or must presume a proposition, sometimes but not always on condition that another proposition has been proved. (To simplify, I will only discuss mandatory presumptions.) For example, it matters for various testamentary purposes whether one person died before or after another person, and yet sometimes it is unclear which it is. Sub-section 184(1) of the Law of Property Act 1925 resolves the impasse by requiring the tribunal of fact to presume that the older person died first. It is not that it has been proved that the older person died first. Indeed, the section applies precisely when there is no evidence that the deaths occurred in a certain order. Instead,

---

<sup>12</sup> They can still draw on everyday experience to reach conclusions from the evidence. See, for example, the Sixth Circuit's *Pattern Criminal Jury Instructions* (n 10) §1.05: 'You should use your common sense in weighing the evidence. Consider it in light of your everyday experience with people and events, and give it whatever weight you believe it deserves. If your experience tells you that certain evidence reasonably leads to a conclusion, you are free to reach that conclusion.'

because there is no evidence, the law requires the judge or jury to proceed on a different basis.

The chart below lists some typical legal presumptions.

---

### Figure 3 – Typical Legal Presumptions

- *That a person accused of a crime is innocent.*
  - *That a child born in wedlock is legitimate.*
  - *That a young child has no criminal intent.*
  - *That a regularly solemnized marriage is valid.*
  - *That a document more than 30 years old is genuine.*
  - *That a person missing for seven or more years is dead.*
  - *That agents are sane.*
  - *That a person intends the natural consequences of his actions.*
  - *That if A buys property from B and directs it to C, C was intended to be a trustee for A.*
- 

What does it mean in law to *presume* a proposition? Thayer's statement is still current:

Presumptions are aids to reasoning and argumentation, which assume the truth of certain matters for the purpose of some given inquiry. ... On whatever basis they rest, they operate in advance of argument or evidence, or irrespective of it, by taking something for granted; by assuming its existence. ... Presumptions are not in themselves either argument or evidence, although for the time being they accomplish the result of both. ... Presumptions, assumptions, taking for granted, are simply so many names for the act or process which aids and shortens inquiry and argument.<sup>13</sup>

Wigmore holds a similar view. This is what he says about the presumption that a person is dead if he has been missing for seven or more years:

... the fact of absence unheard from is to be taken, by a rule of law independent of the jury's belief, as equivalent to death ... [I]f they find the fact of absence for seven years unheard from, and find no explanatory facts to account for it, then by a rule of law they are to take for true the fact of death, and are to reckon upon it accordingly in making up their verdict upon the whole issue.<sup>14</sup>

---

<sup>13</sup> Thayer, *Preliminary Treatise* (n 3) 314.

<sup>14</sup> John Henry Wigmore, *A Treatise on the System of Evidence at Common Law: Including the Statutes and Decisions of all Jurisdictions of the United States* (vol 4, Little, Brown, & Co. 1905) 3533.

Both authors identify similar elements: to presume a proposition is to ‘take for granted’ or ‘take for true’ the proposition when you ‘reason’ or ‘reckon’, regardless of ‘argument or evidence’ or ‘belief’ that it actually is true. Presuming something is a substitute for believing it, or believing the evidence proves it, because it serves the same role in deliberation at trial. It allows you to proceed, deliberately speaking, in a way that might be closed to you if you relied only on what you believed. Modern writers say much the same thing. Edna Ullmann-Margalit, for example, describes presuming something as ‘taking it for true’ or ‘holding it as true’ in order ‘to have a foothold (as it were) for action’.<sup>15</sup> The presumption ‘involves no commitment to, nor guarantee of the truth value of the presumed fact. ... [I]t neither requires nor entitles one to believe [the presumed fact]’.<sup>16</sup>

Legal presumptions have all the features I identified with belief-independent acceptance in the last chapter. First, presumptions lead a judge or jury to ‘take’ or ‘hold’ a proposition as true, or as I would say, to treat it as true. Second, presumptions are context-relative, in that they lead the judge or jury to treat the presumed propositions as true in their trial deliberations, but not in other contexts. Third, presumptions are shaped by practical considerations (a point I return to in a moment). Fourth, the law clearly supposes that presumptions are under the direct control of judges or juries. Finally, presumptions lead judges and juries to treat a proposition as true even if they do not believe it. Indeed, when presumptions have practical relevance it is because they function as a substitute for evidence.<sup>17</sup> For all these reasons presumption is best seen as a type of (belief-independent) acceptance. Section 184(1) of the Law of Property Act 1925, for example, should be

---

<sup>15</sup> Edna Ullmann-Margalit, ‘On Presumption’ (1983) 80 *The Journal of Philosophy* 143, 146, 148

<sup>16</sup> Ullmann-Margalit, ‘On Presumption’ (n 15) 149.

<sup>17</sup> See, for example, Nicholas Rescher, *Presumption and the Practices of Tentative Cognition* (Cambridge University Press 2006) 4: ‘Presumptions by nature provide a provisional surrogate for outright claims to the actual truth.’

understood as requiring that, if it is unclear which of two people died first, the judge or jury is to accept in the context of their deliberations at trial that the older person died first.

The classification of presumption as a type of acceptance has some support in the literature on acceptance. For example, Edna Ullmann-Margalit and Avishai Margalit first ‘equate holding “ $p$ ” as true with having a presumption ... that  $p$ ’, where their understanding of presumption ‘takes its cue from the role of presumptions in the theory of law’.<sup>18</sup> They say that ‘the central intuition behind the use of “accept” ... is taken care of by our notion of “holding as true.” That is, one accepts a sentence when one decides to proceed to act as if it were true, regardless of whether or not one believes it is true’.<sup>19</sup> Putting things together, they should be willing to say that a presumption that  $p$  is acceptance that  $p$ . Similarly, Daniel Mendonca argues that ‘the best candidate for elucidating the idea of presumption as it functions in the legal sphere seems to be the concept of acceptance’, as it is understood by Jonathan Cohen.<sup>20</sup> Cohen himself writes that ‘where a proof depends at any point on a presumption ... a lay trier of fact may be able to accommodate the presumed truth as a reason for accepting the proposed conclusion but be quite incapable of coming to believe that conclusion’.<sup>21</sup> Other writers make similar claims.<sup>22</sup>

---

<sup>18</sup> Edna Ullmann-Margalit and Avishai Margalit, ‘Holding True and Holding as True’ (1992) 92 *Synthese* 167 (1998) 176.

<sup>19</sup> Ullmann-Margalit & Margalit, ‘Holding True and Holding as True’ (n 18) 177-178.

<sup>20</sup> Daniel Mendonca, ‘Presumptions’ (1998) 11 *Ratio Juris* 399, 401.

<sup>21</sup> Jonathan Cohen, *An Essay on Belief and Acceptance* (OUP 1995) 122.

<sup>22</sup> See, for example: Jordi Ferrer Beltrán, ‘Legal Proof and Fact Finders’ Beliefs’ (2006) 12 *Legal Theory* 293, 305; Robert Stalnaker, *Inquiry* (MIT Press 1987) 79: ‘Acceptance ... is a generic propositional attitude with such notions as ... presuming ... falling under it’; Joseph Raz, ‘Reasons:’ in David Sobel and Stephen Wall (eds), *Reasons for Action* (Cambridge University Press 2009) 38: ‘[A]ccepting a proposition is conducting oneself in accord with, and because of, the belief that there is sufficient reason to act on the assumption that the proposition is true ... Acceptance dominates many areas of practical thought. The whole system of law enforcement via courts and tribunals is based on acceptance of presumptions, like the presumption of innocence, and on accepting verdicts based on evidence presented in court, while ignoring all other evidence’.

## *Types of Presumptions*

There are different types of mandatory presumptions. Some are unconditional, in that the judge or jury is required to presume something before any fact has been proved. The presumption of innocence is one example. In general, though, unconditional mandatory presumptions are not called presumptions. Instead the law simply says that the 'burden of proof' is on one party or another. Thus another way to express the presumption of innocence is to say that the prosecution bears the burden of proving the defendant's guilt.<sup>23</sup> Similarly, in a civil case, the plaintiff normally bears the burden of proving the essential facts of the claim. Other presumptions are conditional, in that the judge or jury is required to presume something only upon proof of another fact. For instance, to trigger the presumption that a person is dead must show, among other things, that there is no evidence that the person has been alive for the last seven years.<sup>24</sup>

Presumptions are either rebuttable or irrebuttable (conclusive). That a presumption is irrebuttable means that the judge or jury is required to presume something whatever the evidence to the contrary. Now abolished, the doctrine of constructive malice set a presumption that a defendant who killed a police officer while resisting arrest, or who killed someone while committing a felony, had the intent to kill.<sup>25</sup> Section 76 of the Sexual Offences Act 2003 provides a current example. It says that it must be conclusively presumed that a complainant did not consent to a sexual act if the defendant obtained the complainant's consent by impersonating someone he or she knew. It does not matter that the judge or jury is convinced that the complainant did consent; once it has been proved that the defendant obtained consent by impersonation, they are to proceed as if the complainant did not consent.

---

<sup>23</sup> *Woolmington v DPP* [1935] AC 462 (HL).

<sup>24</sup> *Chard v Chard* [1956] P 259 (DC).

<sup>25</sup> The doctrine was abolished by s 1 of the Homicide Act 1957.

Most presumptions are rebuttable. The judge or jury is to presume some fact unless and until there is sufficient evidence to the contrary. What counts as sufficient evidence depends on the rule, but there are two main categories. An *evidential presumption rule* requires the judge or jury to presume a fact unless and until there is *some* evidence to the contrary. For example, an evidential rule requires the presumption that a person who acted in an official capacity was properly appointed to that office.<sup>26</sup> In contrast, a *persuasive presumption rule* requires the judge or jury to presume a fact unless and until the presumed fact is disproved on a balance of probabilities. The presumption of legitimacy is a traditional example. It is to be presumed that a child born during wedlock is legitimate. To rebut the presumption, it is not enough to introduce *some* evidence that the child is illegitimate, as it would be were the presumption evidential. Rather, it is necessary to prove, on the balance of all the evidence, that the child is illegitimate.<sup>27</sup>

Some presumptions impose a higher standard than even a persuasive standard. The presumption of innocence, say, can only be rebutted by proof beyond a reasonable doubt. In the past, the presumption of legitimacy could likewise only be rebutted by proof beyond a reasonable doubt.<sup>28</sup> There could also be presumptions that set a standard in between the evidential and persuasive standards. We could think of a spectrum, with evidential presumptions rules at one end, persuasive presumption rules in the middle, the presumption of innocence and the old rule regarding legitimacy further along, and so on.

I have mentioned how some presumptions are conditional and others are unconditional, and how some are rebuttable and others are irrebuttable. To anticipate the discussion in the next chapter, unconditional, rebuttable presumptions are analogous to rules

---

<sup>26</sup> *Campbell v Wallsend Slipway & Engineering Co Ltd* (1978) ICR 1015 (DC).

<sup>27</sup> S 26 of the Family Law Reform Act 1969.

<sup>28</sup> Gregory Durston, *Evidence: Text and Materials* (2<sup>nd</sup> edn, OUP 2011) 152. The standard was lowered by s 26 of the Family Law Reform Act 1969.

of thumb. They lead someone to treat something as true unless there is sufficient evidence to the contrary, or sufficient reason to believe the contrary. There is a range of standards, and the degree of certainty or evidence required determines the stringency of the presumption or rule of thumb. By similar reasoning, unconditional presumptions that are irrebuttable are like rules that are not rules of thumb.

### *Functions of Presumptions*

There are three traditional functions of presumptions in law.<sup>29</sup> The first is that a presumption can help to resolve an impasse in deliberation. After all the evidence has been introduced, and the trial enters its final phase, the judge or jury must reach a finding on the disputed facts. They must deliberate, and before long, they must bring their deliberations to a close. If the evidence is conclusive one way or the other, then there is no problem. But what if the evidence is evenly balanced? Presumptions come to the rescue in such situations. They provide a way to bring deliberation to an end, even in the absence of a preponderance of evidence one way or the other. The presumption that the older of two people died first serves this function. Invoking the presumption settles the question, letting the judge or jury reach a finding the evidence alone would not permit.<sup>30</sup>

The second function of presumptions is to reduce the costs of inevitable errors. One way they do that is by mandating a presumption of one fact upon proof of another fact that closely predicts it. Most people inexplicably absent for more than seven years are dead, most children born in wedlock are legitimate, most defendants are sane, etc. Focusing on the

---

<sup>29</sup> See, for example: Edmund Morgan, 'Some Observations Concerning Presumptions' (1931) 44 *Harvard Law Review* 906, 924-931; Richard Eggleston, *Evidence, Proof, and Probability* (2<sup>nd</sup> edn, Wiedenfeld & Nicholson 1978) 108-109; Ullmann-Margalit, 'On Presumption' (n 15) 154-162; Mendonca, 'Presumptions' (n 20) 409-410; Adrian Keane, *The Modern Law of Evidence* (7<sup>th</sup> edn, OUP 2008) 60-62.

<sup>30</sup> Morgan, 'Observations Concerning Presumptions' (n 29) 924-925.

chance of error is enough so long as the costs of each type of error (proceeding on the presumed fact when it does not obtain or vice versa) are equal. But when the costs are unequal, there is an additional question, namely, 'whether one type of error is to be preferred, on grounds of moral values or social goals, over the other(s)'.<sup>31</sup> Ullmann-Margalit writes:

Evaluative considerations may exist which justify a systematic and generic bias in favour of erroneously proceeding on (*q*) rather than erroneously proceeding on not-(*q*), given that (*p*) and given lack of sufficient reasons in the circumstances to believe either *q* or not-*q* to be the case.<sup>32</sup>

What matters is both the chance of a type of error and its costs. Together they determine the acceptability of presuming one fact from another.

There are many examples of social goals moulding presumptions. 'Ownership is inferred from possession because stability of title is desirable'.<sup>33</sup> Children born in wedlock are presumed to be legitimate because at one time illegitimacy was highly stigmatizing.<sup>34</sup> 'The driver of a car is presumed to be driving with the owner's consent in order to give anyone injured by the driver a better chance of recovery'.<sup>35</sup> The best example is the presumption of innocence. Even if the majority of people charged with a crime are factually guilty, the presumption of innocence could still be defended as a way to reduce the costs of errors. Thus Jeremy Bentham writes in his *Treatise on Judicial Evidence* that 'in doubtful cases [the judge should] consider the error which acquits as more justifiable, or less injurious to the good of society, than the error which condemns'.<sup>36</sup> The presumption could fly in the face of probability, but it would be justified because convicting an innocent person is a much worse error than acquitting a guilty person.

---

<sup>31</sup> Ullmann-Margalit, 'On Presumption' (n 15) 159.

<sup>32</sup> Ullmann-Margalit, 'On Presumption' (n 15) 159.

<sup>33</sup> Keane, *Modern Law of Evidence* (n 29) 61.

<sup>34</sup> Thayer, *Preliminary Treatise* (n 3) 346-347.

<sup>35</sup> Keane, *Modern Law of Evidence* (n 29) 61.

<sup>36</sup> Jeremy Bentham, *A Treatise on Judicial Evidence* (JW Paget 1825) 197-198.

The third function of presumptions is to reduce the effort and time the trial takes. For example, the presumption of sanity is justified in part because it would be too time-consuming for the prosecution to have to prove at every trial that the defendant is sane. By allocating the burden of proof, presumptions also ‘tend to prevent the jury’s time being wasted by considering issues unsupported by evidence. They thus shorten the trial by dispensing with the evidence of one party if the other had adduced no evidence to support a finding in his favour’.<sup>37</sup> Still other presumptions aim to place the burden of introducing evidence of a fact with the party for whom it is most convenient to do so. Morgan claims that this is the justification for the presumption that the last carrier of damaged goods is responsible for the damage. ‘Of course there is no balance of probability that it actually was liable for the damage. ... But there is clearly a balance of convenience in its favour in control over the evidence’, the reason being that ‘[e]ach carrier has the opportunity to inspect [the goods] and record their condition upon receipt from the preceding carrier’.<sup>38</sup>

I have mentioned three functions of presumptions: they resolve deliberation, reduce the costs of errors, and save time and effort. I will return to the point in the next chapter, but it should be clear that these are very similar to the traditional reasons for having rules. Like presuming something, having a rule brings deliberation to a close, reduces the costs of errors, and saves people the time and effort involved in going through the weighing process. Of course this may simply be coincidence, or an incidental connection. But if presumption is best thought of in terms of acceptance, and presumption has the same uses as having a rule does, then there is a reason to think rules might be explained in terms of acceptance, too.

---

<sup>37</sup> Keane, *Modern Law of Evidence* (n 29) 61.

<sup>38</sup> Morgan, ‘Some Observations Concerning Presumptions’ (n 29) 926.

*Presumption, Fiction, and Make-Belief*

A legal rule sometimes requires the judge or jury to presume something which they and other interested parties know is false. In that case, the presumption is called a *legal fiction*. The legal rule sets out what is to be presumed, and knowledge of its falsity makes the presumption a fiction. Pierre Olivier, in an exhaustive survey of theories of legal fictions, writes:

A fiction is not a legal rule, but a factual assumption, an acceptance of certain facts contrary to reality. This acceptance can be prescribed by a legal rule, whether statutory or of common law; the meaning of the rule is then an indication to the applier of law that if fact *A* is proven in a certain situation, he must assume, in the place of that fact, that fact *B* had been proven. The legal rule prescribes the fiction, i.e. the factual assumption: the legal rule in itself is not the fiction.<sup>39</sup>

The legal rule tells the judge or jury to ‘accept untrue facts as true and to accept them quite literally, i.e., as if they were the real and proven facts, and to act on these facts’.<sup>40</sup> The fiction is that the untrue facts are the true facts. To engage in the fiction is to proceed on the basis that the untrue facts are true, in other words, to treat the untrue facts as true.<sup>41</sup>

Irrebuttable presumptions are often fictions. When constructive malice was still part of the law, for example, a defendant would sometimes be presumed to have had the intent to kill even though it was clear he lacked that intent. But even rebuttable presumptions may be fictions. That is possible because it could be known the presumed proposition is false, even though the evidence to prove its falsity is inadmissible. For instance, it sometimes arose that after the death of her husband a woman would have a child, and then another a child, within

---

<sup>39</sup> Pierre Olivier, *Legal Fictions in Practice and Legal Science* (Rotterdam University Press 1975) 61.

<sup>40</sup> Olivier, *Legal Fictions* (n 39) 61

<sup>41</sup> The best-known discussion of legal fictions is Lon Fuller’s, published as a series of three articles under the title ‘Legal Fictions’ in the *Illinois Law Review* and then republished as a book. See: (1930-1931) 25 *Illinois Law Review* 363, 513, 877. Fuller claims that ‘a fiction is either, (1) a statement propounded with a complete or partial consciousness of its falsity, or (2) a false statement recognized as having utility’ (369). Fuller’s definition has been criticized many times. Some of the criticisms of the first limb are that for a fiction to be a *legal* fiction it must be required by law; that a fiction in the relevant sense is not a statement; and that the definition is in any event overinclusive because it classes perjury and lies as legal fictions. The second limb is rejected for being both imprecise (recognized as having utility by whom?) and inaccurate (there are legal fictions that at least no longer have social utility). See, for example: Olivier, *Legal Fictions* (n 39) 35; Kenneth Campbell, ‘Fuller on Legal Fictions’ (1983) 2 *Law and Philosophy* 339, 342-345.

the 300-day period that defined wedlock. Obviously the dead husband could not have conceived the second child, but the evidence of his death might not be admissible to rebut the presumption. In that case, the judge or jury would have to presume that the second child was the dead husband's, and thus to engage in a fiction.<sup>42</sup>

Most of the interesting issues around legal fictions are about their significance in the development of the law. For my purposes, that is less important than how natural it is to understand legal fictions in terms of acceptance. Legal fictions lead those who engage in them to treat certain propositions as true despite the knowledge that the proposition is false, just as acceptance is capable of doing. In the last chapter I gave a name to acceptances that do not have corresponding beliefs: 'make-beliefs.' Defined that way, engaging in a legal fiction is a type of make-belief. Make-belief is the broader category, and not just because it includes more than just *legal* fictions. To make-believe a proposition you must *not believe it is true*; to engage in a legal fiction you must *know it is false*.

## 8.5 CONCLUSION

I had two aims in this chapter. The first aim was to demonstrate that acceptance has a place in law. I showed that exclusionary rules of evidence law are in substance requirements *not* to accept certain propositions in a trial context. I also showed that rules requiring presumptions and fictions are in substance requirements *to* accept certain propositions in a trial context. The second aim of the chapter was to lay the groundwork for an analogy between presumptions and rules. I drew connections between the reasons for presumptions and for having rules, and between rebuttable presumptions and rules of thumb. In the next chapter I will also

---

<sup>42</sup> For discussion of a similar case, see: Mendonca (n 20) 409.

bring out their similar motivational roles and in the chapter after, the similarity between a useful legal fiction and an instrumentally justified rule.

# 9 Acceptance and Rule-Possession

---

## 9.1 INTRODUCTION

At the beginning of this thesis, I asked what it is for a person or a society to have a rule. In Chapter 3, I rejected any answer that required a pattern of compliance with the putative rule. That led me to Hart's discussion of the internal aspect of social rules in Chapter 4. Hart showed that having a rule depends on possessing a certain attitude, but he struggled to say what was distinctive about it. To find out more about this attitude, I turned in Chapter 5 to the effect of having a rule on deliberation, and to the relation it bears to certain normative beliefs. The attitude that I was describing seemed, at first, to have much in common with belief in a protected reason, a similarity that I ultimately rejected in Chapter 6. At the end of that chapter, I had settled only that having a rule simulates a belief that the rule ought to be complied with.

In the last two chapters, I gathered the resources to propose an alternate theory of what it is to have a rule. Now, with those resources in hand, I am prepared to set out what I think is a persuasive, acceptance-based theory of what it is to have a rule, which I dub the 'acceptance theory'. Roughly, that theory says that for a person or society to have a rule is for that person or that society's members to accept that the rule ought to be complied with, independent of what they believe. I begin in Section 2 with a discussion of how having a rule and a certain kind of acceptance lead people to act and of how each is related to a belief that the rule ought to be complied with. In Section 3, I discuss some of the other similarities

between the two attitudes. And, finally, in Section 4 I formulate the theory precisely and set out its advantages.

## 9.2 CONTENT, OPERATION, STRENGTH, STRICTNESS

Let me start with three related points about acceptance. First, acceptance is an attitude held towards a proposition, which is the content of the acceptance. Among the propositions you can accept is that an action ought to be performed. Second, acceptance, by its nature, leads you to act in the same ways as a belief with the same content. Acceptance that an action ought to be performed therefore leads you to act as if you believe that action ought to be performed. Finally, acceptance of a proposition is not entailed by a belief in that proposition, and what I termed ‘belief-independent acceptance’ does not entail a matching belief either. These three points are the foundation for an acceptance-based understanding of rule-possession.

I have already said that to have a rule is to have a certain attitude. And I have said that this attitude leads you to act as if you believe that the rule ought to be complied with. This attitude is not entailed by a belief that the rule ought to be complied with, nor does it entail that belief. All of this is equally true of a certain kind of acceptance, namely, acceptance the content of which is that the rule ought to be complied with, where the acceptance is held independent of a belief in that proposition. Here, then, is the crucial point: in the way that they lead a person to act, and in their relation to certain relevant beliefs, having a rule is identical to belief-independent acceptance that the rule ought to be complied with. Put another way, in these ways, having a rule is identical to belief-independent acceptance of the content of the rule or of the rule-proposition.

Take the constitutional rule that the Queen ought to assent to legislation. Constitutional actors, whose rule this is, treat it as settled that the Queen ought to comply with the rule. They say as much, for one thing. They would also exert pressure on the Queen to comply with the rule and criticize suggestions that the Queen should withhold her assent.<sup>1</sup> Yet, despite everything that they do and say, constitutional actors may not *believe* that the Queen ought always to assent to legislation and, even if they did, more would be needed to show that there exists a rule among them that she ought to do so. It all makes for a confusing picture, unless we suppose that constitutional actors *accept* that the Queen ought to comply with that rule, independent of their beliefs on the matter. That allows us to explain the actors' demands and criticisms without attributing to them beliefs that they may not have.

That they are alike in so many ways raises the possibility that what seem to be two attitudes – the attitude in which rule-possession consists and belief-independent acceptance that the rule ought to be complied with – are actually one and the same. I now want to add to the evidence for that possibility by showing that this acceptance-based account of rule-possession offers a simple way to explain the varying strengths and strictnesses of rules.

### *Categories*

There is a distinction between what ought to be done, all things considered, and what ought to be done, relative to reasons of a certain type. Let us then distinguish between acceptance that *As* ought to  $\phi$ , all things considered, and acceptance that *As* ought to  $\phi$ , relative to

---

<sup>1</sup> In 1910 King George V proposed to meet the leader of the opposition in the House of Lords to determine his views on one of the government's bills. Prime Minister Asquith wrote to King George V as follows: 'The part to be played by the Crown, in such a situation as now exists, has happily been settled by the accumulated traditions and unbroken practice of more than 70 years. It is to act upon the advice of the Ministers who for the time being possess the confidence of the House of Commons, whether that advice does or does not conform to the private and personal judgment of the Sovereign'. Quoted in Colin Turpin and Adam Tomkins, *British Government and the Constitution: Text and Materials* (Cambridge University Press 2007) 162-163.

reasons of a certain type. An acceptance may dispose you to treat its content as true in some contexts but not others; it is operative in the first type of context, inoperative in the second. Let us then also distinguish between acceptances that are operative when you *disbelieve* that an  $A$  ought to  $\phi$  and acceptances that are inoperative in such contexts.

Together these two distinctions yield four categories of acceptance.

- (1) Acceptance that  $As$  ought to  $\phi$ , all things considered; operative even when you disbelieve that an  $A$  ought to  $\phi$ , all things considered.
- (2) Acceptance that  $As$  ought to  $\phi$ , all things considered; inoperative when you disbelieve that an  $A$  ought to  $\phi$ , all things considered.
- (3) Acceptance that  $As$  ought to  $\phi$ , relative to a type of reason; operative even when you disbelieve that an  $A$  ought to  $\phi$ , relative to that type of reason.
- (4) Acceptance that  $As$  ought to  $\phi$ , relative to a type of reason; operative even when you disbelieve that an  $A$  ought to  $\phi$ , relative to that type of reason.

My claim is that each category corresponds to a type of rule. Whether the ought is all-things-considered or relative determines whether your rule is strong or weak. Whether the acceptance is operative or inoperative when you disbelieve that the rule ought to be complied with determines whether your rule is strict or a rule of thumb. For example, that you have a strong, strict rule that  $As$  ought to  $\phi$  means that you accept that  $As$  ought to  $\phi$ , all things considered, and that your acceptance is operative even when you disbelieve that an  $A$  ought to now  $\phi$ , all things considered.

This chart shows the four possibilities. The reference to ‘ought’ in the second column should be read as referring to an all-things-considered or relative ought, as appropriate.

<b>Figure 4 – Types of Rules</b>		<b>Content of Acceptance</b>	
		<i>As ought to <math>\phi</math>, all things considered</i>	<i>As ought to <math>\phi</math>, relative to a type of reason</i>
<b>Operation of Acceptance</b>	<i>Operative when you disbelieve an A ought to <math>\phi</math></i>	(1) Strong/Strict	(3) Weak/Strict
	<i>Inoperative when you disbelieve an A ought to <math>\phi</math></i>	(2) Strong/Rule of thumb	(4) Weak/Rule of thumb

This chart shows how various kinds of acceptances map on to the possession of various kinds of rules. What all these acceptances have in common is that they are acceptances that a rule ought to be complied with. That much rule-possession always shares, on the acceptance-based account. But it is the differences that matter for the moment, and in the rest of this section I give some examples designed to demonstrate these connections.

### *Content and Strength*

In a classic legal anthropological study, Margaret Hasluck describes the customary law in Albania in the early 20<sup>th</sup> century, including what she called ‘the law of the dog’.<sup>2</sup> Because dogs were crucial for security and work in rural Albania at that time, the law governing their treatment was extensive. Different laws applied depending on the type of dog. One

---

<sup>2</sup> Margaret Hasluck, *The Unwritten Law in Albania* (Cambridge University Press 1954) ch 7.

particularly important type of dog was called a ‘chained dog’. ‘Chained dogs were ... dogs with the function of guarding their master’s house from thieves and enemies. The poorest man kept one and regarded it as a valuable article of property’.<sup>3</sup> To make them fiercer, chained dogs were never shown affection or allowed indoors. But, as a kind of compensation for their harsh treatment, the law of the dog prohibited anyone from killing a chained dog at night, when they performed their duties.<sup>4</sup>

The prohibition against killing a chained dog at night was exceptionally severe, as Hasluck explains.

During the night the law was entirely on the side of the house dog [including a chained dog] and its master. No matter how the animal might threaten a stranger he might not kill it. ... The Kanùn [a codification of legal rules] says clearly that ‘at night a dog is equivalent to a man.’ Indeed, in some places such as Çermenikë the animal was thought more important and its murder by night was more serious.<sup>5</sup>

Killing a chained dog at night was treated as unjustified even when the killing was done in self-defence and in response to an unprovoked attack. In Hasluck’s telling, one night in the region of Lure, a woman went out in her own farmyard, and there ‘[a] neighbour’s dog sprang on her and killed her, but since it was the night time, the dog’s time, her blood was lost. It was immaterial that the dog was trespassing and that she was on her own ground’.<sup>6</sup>

It seems reasonable to say that rural Albanians, at least in Çermenikë and Lure, treated the killing of a chained dog at night as unjustified whatever the reasons to the contrary. They thus had a *strong* rule against killing a chained dog at night. The strength of their rule can be explained, on an acceptance-based account of rule-possession, through the content of

---

<sup>3</sup> Hasluck, *The Unwritten Law* (n 2) 73.

<sup>4</sup> Hasluck describes one exception: ‘If a house stood near a main road an aggressive dog might be killed during the night without a penalty. A higher law stated that “it is not right to cut a main road with a dog.”’ *The Unwritten Law* (n 2) 75.

<sup>5</sup> Hasluck, *The Unwritten Law* (n 2) 75.

<sup>6</sup> Hasluck, *The Unwritten Law* (n 2) 75.

their acceptance. Had they accepted that a chained dog ought not to be killed at night *all things considered*, Albanians would have accepted that there is a reason not to kill a chained dog that is as weighty as any contrary reason. No reason would outweigh it; nothing would override it. Having accepted such a strong reason, Albanians would have condemned any killing of a chained dog at night, no matter what the cost. They would have behaved, in other words, exactly as the inhabitants of Lure did in Hasluck's story.

The law of the dog differed from one community to the next, and in some communities some of its rules were weak rules. That is clearest from Hasluck's description of the law governing sheep dogs. In the poor north of Albania, 'where men depended entirely on their flocks, the sheep dog was of paramount importance'. The law there granted sheep dogs similar protection to that afforded chained dogs.

Among the mountains of the north the law of the sheep dog was the same by day and by night. A passer-by might be bitten at any hour of the day or night, but if he killed the dog he would be killed by its master. No excuse for his crime could be tolerated. He might have lost his way or be a stranger (and so a guest to be protected). No matter! The mountain was the dog's ground day and night.<sup>7</sup>

As in the case of chained dogs, the law in the north protected sheep dogs within a certain sphere. Chained dogs were granted protection during the night, but not the day. Sheep dogs were granted protection day and night, but only on the mountain near the flocks. In both cases the dog was protected in the sphere in which it performed its duties. But the law in the north did not extend to the more prosperous south. 'In fertile Lumë', Hasluck writes, 'the law was that "a man can't lose his life for a dog", and on that theory a traveller on the mountain might kill a sheep dog which had refused to let him pass'.<sup>8</sup>

---

<sup>7</sup> Hasluck, *The Unwritten Law* (n 2) 77-78.

<sup>8</sup> Hasluck, *The Unwritten Law* (n 2) 78.

Communities in both the north and the south had a rule against killing a sheep dog, but the rule was a strong rule in the north and a weak rule in the south. The difference lies, on the acceptance-based account, in the kind of ‘ought’ the members of these communities accepted. The inhabitants of Lumë would have accepted that a sheep dog ought not to be killed, relative to non-survival-related reasons, which would explain why they allowed that a traveler may kill a menacing sheep dog. The northerners, on the other hand, treated such killings as unjustified, which would be explained by their acceptance that a sheep dog on the mountain ought not to be killed all things considered. Put another way, the Lumëans accepted a reason not to kill sheep dogs that may only be overridden by survival-related reasons, whereas the northerners accepted a reason that is not overridden even by survival-related reasons.

Here is an analogy that might help. In the last chapter, I discussed presumptions in the law of evidence and how they can be understood as acceptances. Some presumptions, left unrebutted, resolve the *overall* issue at trial.<sup>9</sup> The presumption of innocence is one example. These presumptions are like strong rules, which resolve what ought to be done, all things considered. Thought of this way, northern Albanians presume that the answer is ‘no’ to the overall issue of whether a sheep dog on a mountain ought to be killed. Other presumptions resolve a subsidiary issue at trial, such as whether a defendant had the intent to kill. These presumptions leave room for evidence and argument, because the overall issue still needs to be determined. Similarly, the Lumëans presume that a sheep dog on a mountain ought not to be killed, relative to *certain* reasons, but they are open to persuasion on the overall issue.

In summary, on an acceptance-based account, the difference between having a strong versus a weak rule is the difference between acceptance that the rule ought to be complied in

---

<sup>9</sup> It is a simplification to suppose that there is always one overall issue at trial, or that there is a clean distinction between an overall and a subsidiary issue. If you prefer, you can assume I am only talking about presumptions in trials at which there *is* a single overall issue, there *is* a distinction between that issue and subsidiary issues, etc.

an all-things-considered or relative sense. This way of thinking makes sense of the law of the dog. And it can make sense of my earlier examples, too. One example was of a rule among gentlemen of the American South that required them to accept a challenge to a duel. That rule led gentlemen to risk death, but it yielded when necessary to save their honour. On an acceptance-based account, gentlemen would have accepted that they ought to agree to a duel, relative to all non-honour-related reasons, including reasons of survival.

### *Operation and Strictness*

In 1959, the Canadian Department of Fisheries published the *Canadian Fish Cook Book* in an effort to increase fish consumption in Canada.<sup>10</sup> The book gives concise instructions for scaling, gutting, skinning, boning, and especially cooking fish. The basic guidance is clear and commonsensical:

You will know your fish is cooked when:

1. The flesh loses its translucent appearance and becomes opaque.
  2. The flesh flakes readily.
  3. The flesh is easily pierced with a fork.
- ...

Remember that any fish can be cooked by this method. ... [Y]ou will note that some recipes call for a certain variety of fish. You may substitute other varieties of fish and still obtain good results.<sup>11</sup>

The book is a classic among cooks because it also sets out what is sometimes called the ‘Canadian Rule’: Cook a (thawed) fish for 10 minutes per inch measured at its thickest point. The Department of Fisheries is very clear about this: stuffed fish ought to be cooked for 10 mins/inch, as should fish that is not stuffed; fish should be broiled, steamed, deep fried, or

---

<sup>10</sup> Canada Department of Fisheries and Forestry, *Canadian Fish Cook Book* (Canada Department of Fisheries and Forestry 1959) 8.

<sup>11</sup> Department of Fisheries, *Canadian Fish Cook Book* (n 10) 8.

pan fried for 10 mins/inch; so should fish cooked in liquid. The only variation allowed is in the cooking temperature. Generally it should be 400°F, but it depends on the type of fish.

The Canadian Rule is a helpful guide to how to cook fish, to be set aside when a cook is confident enough that a fish would be tastier cooked another way. It is, in short, a rule of thumb. Because the Canadian rule is ‘nearly foolproof’<sup>12</sup>, it will not be set aside very often. Amateur cooks, in particular, may require the near certainty of a tastier fish to feel comfortable departing from the rule. Expert chefs will likely feel differently. They will know, for instance, that fish cooked by super hot searing needs less time, so they will demand less certainty from themselves to depart from the rule, and will emphasize other indicators instead (such as 1 to 3 in the list above).<sup>13</sup> In either case, there will be situations in which the cook is inclined to simply treat it as true that a fish should be cooked for 10 mins/inch, and situations in which he or she will instead act on his or her judgments of the merits of cooking a fish for one length of time or another.

The explanation for the cooks’ behaviour lies, I claim, in their possession of acceptances with a certain scope of operation. Cooks who take the Canadian Rule as a rule of thumb accept that a fish ought to be cooked for 10 mins/inch, relative to reasons of taste and pleasure and the like, but their acceptances are inoperative when they are sufficiently confident that *this* fish would be tastier if cooked differently. In such contexts, their acceptances cease to motivate them, and they naturally fall back on their beliefs about how best to cook the fish. For amateur cooks, that confidence level is set rather high; for expert cooks, it is set lower. As a result, the scope of operation of an amateur cook’s acceptance is wider than that of an expert cook.

---

<sup>12</sup> Marian Burros, ‘Plain and Simple; A Dish So Good Even the Fat is O.K.’ *New York Times* (New York, 7 April 1993) <<http://www.nytimes.com/1993/04/07/garden/plain-and-simple-a-dish-so-good-even-the-fat-is-ok.html>> accessed 11 July 2011.

<sup>13</sup> See, for example: John Taylor, *Hoppin’ John’s Low Country Cooking: Recipes and Ruminations from Charleston and the Caroline Coastal Plain* (Houghton Mifflin Harcourt 2000) 73-74.

Rules of thumb, on this analysis, are analogous to rebuttable presumptions in the law of evidence. Rebuttable presumptions serve as provisional substitutes for evidence of the presumed proposition. In the same way, rules of thumb take the place of a belief in what ought to be done. Rebuttable presumptions yield in the face of sufficient, admissible evidence to the contrary, and rules of thumb yield to a contrary belief of sufficient confidence. When a cook simply proceeds on the basis that a fish should be cooked for 10 minutes per inch, it is like when a judge simply proceeds on the basis that the older of two people died first in a simultaneous accident. The cook does not rely on his or her judgment, and the judge does not rely on evidence.

As rules of thumb are to rebuttable presumptions, strict rules are to irrebuttable presumptions. Irrebuttable presumptions lead someone to treat the presumed proposition as true no matter what the evidence to the contrary. They are not a *provisional* substitute for evidence; they are a *final* substitute. Likewise, when someone has a strict rule, he or she treats it as true that the rule ought to be complied with, regardless whether it is clear that is not the case. The explanation, on the acceptance-based account, is that the person's acceptance is operative regardless whether he or she disbelieves that the rule ought then to be complied with.

When there is a rebuttable presumption at trial, one party may seek to rebut the presumption by admitting evidence to the contrary. Similarly, if we each take it as a rule of thumb that a fish should be cooked for 10 mins/inch, but you are trying to persuade me to cook this fish for longer, you will try to show that the case for cooking this fish for more time is especially clear or obvious. You will try, in other words, to raise my level of confidence that the fish would taste best if cooked for more than 10 mins/inch. Strict rules lead to a different dynamic. If there is a strict rule of etiquette that a fish must be eaten with a fork, unless silver knives are provided, but we disagree about whether to use a knife for a particular dish, our

argument will likely be about whether the rule applies. (Is this a fish? Is this knife silver?) What we would *not* do is try to resolve this practical matter by investigating all the merits of eating fish with a knife. Likewise, when a presumption is irrebuttable, it is no good trying to admit evidence against the presumed proposition. The party resisting the presumption must instead admit evidence against the fact that triggered the presumption in the first place.<sup>14</sup>

In the last section, I analogized strong rules to presumptions on the overall issue at trial and weak rules to presumptions on subsidiary issues. Now I have analogized strict rules to irrebuttable presumptions and rules of thumb to rebuttable presumptions. I should make clear, however, that in the law of evidence there are no irrebuttable presumptions on the overall issue. Otherwise, there would be no need for a trial. (Just think of what it would mean for there to be an irrebuttable presumption of innocence.) This is where the analogy breaks down, because some societal rules *are* both strict and strong, including, for example, the northern Albanians' rule against killing a sheep dog on the mountain.

### *Communication and Sincerity*

Throughout this section, I have said that the acceptance that a rule ought to be complied with leads people to act as if they believe that the rule ought to be complied with. I have said, more specifically, that the acceptance leads people to 'criticize' violations from the rule and to 'condemn' those who violate it; to 'pressure' others to comply with the rule and to 'praise' them when they do. It leads them, in short, to make the sort of communications that Hart

---

<sup>14</sup> This is also a strategy a party resisting a *rebuttable* presumption could use. There are thus *two* options open to a party resisting a rebuttable presumption: admit evidence that tends to disprove the triggering fact, admit evidence that tends to disprove the presumed fact. And that makes sense, on my analysis. I could try to convince you not to cook our fish for 10 minutes/inch by showing that the rule does not apply – we are making sushi, for instance.

showed were associated with the existence of a rule in a society. And that, I have claimed, is one of the main advantages of an acceptance-based account of rule-possession.

But I have snuck in an assumption that you may not like. The acceptances I have been referring to are belief-independent acceptances, and so I have allowed that people may have a rule without believing that there is a reason to comply with it. That means I have assumed that people may criticize and praise others without actually believing that what they did was justified or unjustified, as the case may be. You might think, though, that true criticism or praise is always sincere, and sincerity depends on having just these kinds of beliefs. What I have described is merely the pretence of having a rule, not its reality.

Let me grant for the sake of argument that belief is essential for sincerity, and that sincerity is essential for criticism or praise. Then the key point is how sure we are that those who have a rule *actually* criticize its violation rather than *appear* to do so. If someone says, ‘You should have worn black to the funeral’, it will normally be the case that he or she believes that. Yet assuming what is normally the case is actually the case for a particular statement is unjustified if there is another plausible explanation for it, as there is for the communications made by people who have a rule, on an acceptance-based account. It then becomes a straight assumption to say that people who have a rule sincerely criticize its violation.

There is a supporting point. Although belief-independent acceptance does not imply a belief with the same content, it does not imply the lack of that belief either. Indeed, sometimes such an acceptance leads to a matching belief. Suppose, then, that the members of a society generally accept that a rule ought to be complied with, and this eventually leads them to believe the same. What were simulated criticisms will become sincere criticisms. If this happened often enough, or if belief and acceptance coincided for some other reason, it might give the impression that a disposition to sincerely criticize the violation of a rule is

necessary for its existence in a society. In fact, though, the two would just tend to go together. Another consideration is that people sometimes have instrumental reasons to comply with their rules, which may lead them to believe that these rules ought to be complied with (I discuss the point in detail in the next chapter).

### *Protected Reasons*

In Chapter 6, I proposed, for the sake of argument, a protected reasons-based account of rule-possession. One of the unattractive features of that account was that it was limited to strict rules; it could not explain rules of thumb. I did not dismiss the protected reasons-based account on that basis, however, partly because it was not yet apparent whether there is a unified account on offer. Now it is clear that there is. Although I have yet to fully articulate an acceptance-based account of rule-possession, it should be evident that we can use acceptances to explain the way that people act when they have a strict rule or rule of thumb, as well as a strong rule or a weak rule. That is an advantage of an acceptance-based account, relative to the protected reasons-based account.

Another unattractive feature of the protected reasons-based account, I claimed, is that it requires that exclusionary reasons take priority over the reasons they exclude, and we have no grounds for thinking that they do. Since rejecting the protected reasons-based account, I have proceeded on the basis that, in fact, the balance of reasons determines what ought to be done. That has made a significant difference to my argument. It made it possible, at the end of Chapter 6, to claim not merely that having a rule does not entail the belief that the balance of reasons favours compliance, but that it does not entail the belief that the rule ought to be complied with. And it is why, in this chapter, I claimed that rule-possession is the *belief-independent* acceptance that the rule ought to be complied with. So, the earlier discussion of

exclusionary reasons matters, not just to whether we should reject the protected reasons-based account, but for what counts as a good acceptance-based account.

What if I was wrong and I should have proceeded on the basis that exclusionary reasons *do* take priority over the reasons they exclude? In that case, the acceptance-based account would have to be modified. I could no longer claim that having a rule does not entail the belief that it ought to be complied with. I would have no basis for focusing on acceptances that are belief-independent. That is all, though. One of the more interesting features of the account would be lost, but its core would remain the same. According to the alternate account, having a rule is like the acceptance that it ought to be complied with, and it is left open whether that acceptance is belief-independent or belief-dependent.

### *Summary*

So far I have argued for four points. First, the possession of a rule and the belief-independent acceptance that it ought to be complied with lead people to act in the same way. Second, such an acceptance neither entails nor is entailed by a belief that the rule ought to be complied with, features which it shares with possession of that rule. Third, by varying the content or the operation of the acceptance, it is possible to account for the ways people are differently disposed when they have rules of varying strengths and strictnesses. The last point was that rule-possession is analogous to making a presumption, and presumptions, as I have already said, can be thought of as acceptances. These points all support an acceptance-based account of rule-possession, and in the next section I make additional arguments in its favour.

### 9.3 REASONS, VOLUNTARINESS, AND CONTEXT-DEPENDENCE

My focus in the last section was on the way an acceptance disposes a person to act and on how it is related to some of this person's beliefs. Acceptance has other features, though, and the task now is to see whether they, too, have their counterparts in rule-possession.

#### *Practical Reasons*

Acceptance, unlike belief, is shaped by practical reasons. You can accept a proposition for practical reasons, in a way that you cannot acquire a belief for practical reasons. The rules that people have are shaped by practical reasons, too. Indeed, the traditional reasons for having rules are practical, instrumental reasons. As I discussed in Chapter 5, people often have rules that simplify their deliberations and to help them avoid error. Moreover, they have such rules *because* having them serves their goals. Recall Jill's case. It is stressful for Jill to weigh all the merits and demerits various holiday destinations, year after year. It is wearying to have to determine dates and calculate costs. Having a rule to spend her holidays in France simplifies Jill's deliberations, and she acquires the rule for that reason. In Raz's telling, '[Jill] adopted a rule to spend her holidays in France and she did so precisely in order to spare herself the necessity of deciding every year what to do during the holidays'.<sup>15</sup> I give several examples that demonstrate the same point in the next chapter.

Acceptance is not just responsive to practical reasons; it is *un*responsive to evidence for and against what is accepted. Suppose that Jill is preparing to book her holidays when she sees an advertisement for a cheap flight to Spain. To Jill it seems plain that, this year, she should go to Spain. Will this new evidence lead Jill to revise her rule? Frederick Schauer

---

<sup>15</sup> Joseph Raz, 'Reasons for Action, Decision, and Norms' (1975) 84 *Mind* 481, 497-498.

would say ‘no’. ‘[R]ules may [not] be subject to exceptions at the moment of application in the full range of otherwise applicable factors and still be rules’.<sup>16</sup> And this makes perfect sense on an acceptance-based account of rule-possession. Accepting that she ought to go to France on holidays, Jill will not be tempted to abandon that acceptance just because the evidence against that proposition has strengthened. Indeed, if Jill continually updated what she ‘accepted’ to reflect the evidence, we would not say she had an acceptance at all – or a rule, according to Schauer.

Someone might object that the cheap flight is not merely evidence that what Jill accepts is false. It is a practical reason for her to revise her acceptance, because the acceptance gets in the way of her obtaining a cost savings. Because acceptance is responsive to practical reasons, we should expect Jill to be open to revising her acceptance in response to the cheap flight. But suppose that Jill really did revise her acceptance on this basis. She would have revised her rule because she believes that she ought to go to Spain and revising her rule is a step towards going there. That is not something she would do if she truly believed that she ought to go to France. So, in revising her rule, Jill has not proceeded as if she believes she ought to go to France, which only makes sense if she has *already* ceased to accept that she ought to go to France. The objection is circular: it seeks to *explain* why Jill would abandon her acceptance, but in fact it *presumes* that she has abandoned it.

This objection brings out an interesting feature of acceptance. Jill’s acceptance – and all similar acceptances – is what we might call ‘self-protective’: the acceptance protects itself against revision motivated by the possessor’s belief that revising the acceptance will lead him or her to do as he or she ought not to do. When it comes to rules, it is more natural to say

---

<sup>16</sup> Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (OUP 1991) 116. Schauer thinks that adding an exception to a rule is the same as replacing the first rule with another rule: Frederick Schauer, ‘Exceptions’ (1991) 58 *University of Chicago Law Review* 871, 876. See also his ‘On the Supposed Defeasibility of Legal Rules’ (1998) 51 *Current Legal Problems* 223, 226-227. Whether or not Schauer is correct about this is not important for my purposes. If you prefer, you can read the passage in the main text as being about rule-revision, rather than about the addition of exceptions to rules.

that you will not revise your rule just because, in the circumstances, it strikes you as overinclusive.

I should caution that Jill is not simply stuck with her acceptance, or her rule. For one thing, Jill is unlikely to accept that she ought to go to France *all things considered*. She will accept that she ought to go there relative to reasons of price and convenience and the like. Other types of considerations – safety-related reasons, for instance – may lead Jill to take her holiday elsewhere, or to revise her rule. She may also lose her acceptance through some non-deliberate, tacit process. There are also considerations that are relevant to what Jill ought to accept, but which do not bear on where Jill ought to take her holiday. Perhaps Jill does not want to be the sort of person who makes decisions about where to go on holiday based on a rule. Or perhaps she has decided that debating the merits of various holiday destinations was actually rather fun, and she would like to go back to doing so. These are just a few examples; there are other possibilities.

### *Voluntariness*

Acceptance, unlike belief, is under a person's direct, voluntary control. Without using any means, you can make it the case that you accept or do not accept a proposition. In contrast, you cannot make yourself believe something without using some means of doing so, such as hypnosis or selective exposure to evidence. Whether a society has a rule turns on whether a sufficient number of its members have a certain attitude. For that reason, is not under any *one* person's control whether the society has that rule. But whether to have a *personal* rule is a different matter. It is possible to acquire or abandon a personal rule without using the indirect methods necessary to acquire or lose a belief. Jill, in Raz's example, can bring it about that she has a rule to go to France on holiday, or does not have that rule, at will. To be clear, I do not

mean that *every* personal rule is the product of a decision of the person whose rule it is. It is possible to acquire a personal rule tacitly, just as one may come to accept something tacitly.

This is a good time to mention one of the possible implications of the acceptance-based account. It is an account of rule-possession. It is not an account of what it is to enact or make a rule. Yet the two topics may turn out to be more closely connected than they seemed at first. For if rule-possession consists in acceptance that a rule ought to be complied with, and you can voluntarily make it the case that you accept something, then you can voluntarily make it the case that you have a rule. You can, in other words, choose your personal rules. Now, 'voluntarily making it the case that you have a rule' sounds very much like 'making a rule for yourself'. This is, I think, the way we would describe what Jill does: she makes a rule for herself to go to France on holiday. I have not been discussing how to make a rule, so I do not want to oversimplify things, but there is at least the possibility that, at the level of personal rules, making a rule may be a matter of performing the mental action that results in the right sort of acceptance.

Can the same idea be extended beyond personal rules? Suppose that you perform an action, such as making a statement, with the intention of someone else thereby accepting that some action ought to be performed, independent of his or her beliefs on the matter. And suppose that, as a result of your action, that other person does, in fact, acquire that acceptance. Then, if I am right in thinking that rule-possession consists in this kind of acceptance, we can say that you made it the case that someone had a rule. Hence you made it the case that there *is* that rule. Would we also say, you made that rule? I do not have a clear sense of the answer to this question. Perhaps more is necessary to make a rule, and certainly there are other ways to make a rule. This line of thinking does, however, suggest that there may be a way to extend the acceptance-based account of rule-possession to interpersonal rule-making.

### *Hart and Acceptance*

Recall from chapters 3 and 4 that, according to Hart, a society has a rule just if there is a pattern of conduct in that society and enough of its members possess a certain ‘critical reflective attitude’. Hart is often read as describing that attitude only in terms of its manifestations, which is what makes Warnock’s Objection so powerful. That objection is that Hart has not told us what distinguishes the attitude in question from a belief in a reason to comply with the rule, which makes his theory overinclusive. There is, however, another way to read Hart, one that I have mentioned already.<sup>17</sup> Hart sometimes calls this critical reflective attitude ‘acceptance of a rule’. Under this name, he describes the attitude as having some of the same or similar features as the kind of acceptance that I have cast as the key to rule-possession. I am thinking particularly of his claim that acceptance of a rule is an attitude that is responsive to a variety of reasons, including practical reasons, and that the attitude may be voluntarily acquired. Belief, as I have just said, does not have these features, so it may be that there were the resources in Hart’s theory to respond to Warnock’s Objection.

At first it seems that Hartian acceptance, held as it is towards a rule, must be different from the kind of acceptance I am talking about, which is held towards a proposition. But that may not be the case. Suppose you were asked: what is a rule? There are different kinds of answers that might be fitting. You could say what sort of thing a rule is, namely, an abstract object of some description. Or, you could give an example of a rule by stating the normative proposition that is the content of a rule. You could say, for instance, ‘forks ought to be placed to the left of the plate – that is a rule’. There are parallel ways of reading the phrase ‘acceptance of a rule’. On one reading, acceptance of a rule is acceptance of the object that is

---

<sup>17</sup> See Section 4.5; also 7.3.

a rule. Understood this way, acceptance of a rule is like acceptance of a situation or a bribe; it is *not* acceptance of a proposition. On another reading, acceptance of a rule is acceptance of the normative proposition that is the content of a rule. So long as we read Hart in this second way, my account is not obviously inconsistent with his account, at least in this respect.<sup>18</sup> That our society has a rule that forks ought to be placed to the left of the plate implies, on either account, that enough of us accept that rule, i.e., accept that forks ought to be placed to the left.<sup>19</sup>

Our society does not have a rule that forks ought to be placed to the left of the plate *merely* because we accept that forks ought to be placed to the left of the plate. Our acceptances must also be belief-independent, for one thing. And there is another point, one that is suggested by Hart. Our acceptances must operate in at least some contexts in which it is of *practical* relevance whether forks ought to be placed to the left. It is easiest to see why through another example. In a lecture on Roman law, students may be asked to suppose, and hence accept, that slaves may not vote.<sup>20</sup> Even if they all supposed that, we would not say that the students had a rule against slaves voting, because their acceptances would be used only to draw theoretical conclusions, not to guide their actions. Nor, interestingly, would we say that the students accepted *a rule* that prohibited slaves from voting. Perhaps this is no coincidence. It could be that Hart used the phrase ‘acceptance of a rule’ to mark the difference between mere acceptance of the content of a rule, on the one hand, and acceptance that has the further features necessary for the existence of that rule, on the other. On my account, that

---

<sup>18</sup> There is a reason for adopting the second reading over the first reading of Hart. For Hart, the *existence* of a social rule depends on the members of a society possessing the appropriate critical reflective attitude, or, in other words, accepting that rule. If acceptance of a rule means acceptance of the object that is that rule, then a rule must first exist before someone can accept it. Hart would then have made the existence of a social rule a condition of its own existence – a problematic result. No such circularity arises on the second reading of Hart.

<sup>19</sup> Strictly speaking, what is necessary is that enough members of a society accept the *putative* rule.

<sup>20</sup> Cf. Hart, *The Concept of Law* (2<sup>nd</sup> edn, OUP 1994) 105: ‘One vivid way of teaching Roman Law is to speak *as if* the system were efficacious still and to discuss the validity of particular rules and solve problems in their terms ...’.

would mean that 'acceptance of a rule' stands for practical, belief-independent acceptance of the content of that rule.

I should be clear that my claim is only that there is a way to read Hart according to which my account is compatible with his, and perhaps even continuous with it. I do not claim that this is the most natural or faithful reading of Hart. However, if the account I argue for is attractive, it would be a charitable reading of Hart. Nor do I claim that the contemporary understanding of acceptance in the philosophy of action is foreshadowed in its entirety in Hart's discussion of the acceptance of a rule. It is not, and that is clearest when it comes to the context-dependence of acceptance.

### *Context-Dependence*

Acceptance is context-dependent, meaning it leads you to treat its content as true in one context but not another. The contexts in which the acceptance so disposes you are the contexts in which it is operative. Rule-possession is context-dependent, too. We know that because having a rule of thumb disposes you differently depending on what you believe ought to be done in the context. The context-dependency of rule-possession can also help explain why people who have a rule because they occupy a role often act so differently in role- and non-role-based contexts.

Consider a judge who takes it as a rule that hearsay ought to be excluded. When in court, he says things like, 'Hearsay ought to be excluded'. He rules against the admission of hearsay evidence. He gives directions to the jury warning them against using hearsay in their deliberation. He makes critical statements to counsel when they attempt to introduce hearsay. So long as he is in court, the judge acts in all respects as if he was firmly convinced that hearsay should be excluded. Out of court, though, the judge behaves very differently. He

openly criticizes the exclusionary rule. Hearsay, he says, should always be admitted. Indeed, he is a vocal advocate for a change in the law. He expresses these sentiments vigorously, and seemingly sincerely. When he returns to court, though, the judge again speaks and acts as if he believes hearsay should be excluded.

The judge's behaviour would be hard to explain if we thought he believed everything he said. The judge would be shifting back and forth between the belief that hearsay should always be excluded and the belief that it should always be admitted. That would be irrational, absent changing evidence. It is also implausible, as a general explanation, given how frequently judges criticize the rules they apply. On an acceptance-based account of rule-possession, in contrast, the judge's behaviour has a natural explanation. He accepts that hearsay ought to be excluded and his acceptance is operative in court and inoperative out of court. His acceptance leads the judge to act as if he believes hearsay ought to be excluded when he is in court, but not when he is out of court. His behaviour is explained, not by a change in attitude, but by a change in context.

To take another example, suppose a dean of admissions has it as a rule that all men and women should be treated equally.<sup>21</sup> When acting in his official capacity, he is impartial and even-handed. He says things like, 'it would be wrong to treat anyone differently because of their sex' and 'everyone deserves equal consideration, regardless of race'. He tells others that they are behaving objectionably when they make sexist or racist comments or jokes, and urges others to examine their views for signs of prejudice. And yet the dean does not really believe all men and women should be treated equally. Secretly he harbours racist and sexist views. In private, the dean is true to his beliefs, and makes all kinds of prejudiced statements. On an acceptance-based account, the dean's behaviour is easily understood if he accepts that

---

<sup>21</sup> For a similar example, see: Edna Ullmann-Margalit and Avishai Margalit, 'Holding True and Holding as True' (1992) 92 *Synthese* 167, 178.

all men and women ought be treated equally and his acceptance is operative in official contexts and inoperative in private contexts. That acceptance would lead him to act as if he believes all men and women should be treated equally when the acceptance is operative, which is why the dean makes tolerant-sounding statements at work. In his private capacity, the acceptance is inoperative, and the dean proceeds on the basis of his racist and sexist beliefs.

### *The Anarchist Judge*

So far I have discussed how rule-possession and acceptance are alike in being responsive to practical reasons, under a person's direct and voluntary control, and context-dependent. I have discussed these features separately, however. Here is an example of how they come together.

Recall that a rule of recognition is a social rule, one that requires officials of the system of which they are part to apply the rules it identifies as valid. The existence of the rule of recognition depends on the attitudes of these officials. It depends, Hart would say, on their acceptance of the rule. The officials can possess this attitude for a variety of reasons, including practical reasons. Raz provides the useful example of an anarchist judge, who reasons that 'if he follows the law most of the time he will be able to disobey it on the few but important occasions when to do so will tend most to undermine it'.<sup>22</sup>

Raz does not describe the anarchist judge in detail, so let me imagine the rest of the story. Most of the time, the anarchist judge will act as if he believes he ought to apply the laws

---

<sup>22</sup> Joseph Raz, *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990) 148. Cf. Joseph Raz, *The Authority of Law: Essays on Law and Morality* (OUP 1979) 156 n 13: 'Hart maintains that judges at least weakly accept the rule of recognition [i.e., believe that they ought to follow it themselves]. A judge who merely weakly accepts it must, it would seem, pretend that he fully endorses it'. Hart comments on these passages of Raz's in his *Essays on Bentham* (OUP 1982) 155-159.

of the legal system of which he is part. He will act that way in court, say, by applying the laws. He will act that way when making public appearances. And he will do all this to further his long-term goal of bringing down the legal system. In other contexts, though, the judge's true character will shine through. When he attends meetings of an anarchist society, or plots strategy with other anarchist judges, he will express his belief that no one – himself included – ought to apply the law. The judge will lead a double life. In public he will treat as true what he does not really believe; only in private will the two coincide.

The judge is like an actor in a play. Just as an actor may take the point of view of his character for greater verisimilitude, the judge chooses to accept the rule of recognition to advance his anarchist ends. When the play is over, the actor will shrug off his persona, simply by deciding to do so. Likewise, when the time is right, and he determines that not applying the law will maximally damage the legal system, the judge will cast off his acceptance and act from conviction. It is up to the actor to choose when to play his part and when not to, and it is up to the judge to decide whether to accept the rule of recognition and for how long.

The anarchist judge accepts that he ought to apply the law, and he does so as a matter of choice, for a practical reason, and relative to some but not all contexts.

### *Summary*

In this section, I added to the argument for an acceptance-based account of rule-possession. I claimed that rule-possession shares several important features with acceptance: they are both responsive to practical reasons, under a person's direct and voluntary control, and context-dependent. To show that rule-possession has these features, I drew on examples of personal rules, role-based rules, and rules held for instrumental reasons.

#### 9.4 THE ACCEPTANCE THEORY

In Section 2, I claimed that having a rule leads people to act in the same way as acceptance that it ought to be complied with. If this acceptance is belief-independent, then neither attitude entails nor is entailed by a belief in that reason. In Section 3, I claimed that acceptance and rule-possession are each responsive to practical reasons, susceptible to direct and voluntary control, and context-dependent. I also made clear that an acceptance that leads someone to act in the ways that having a rule does must be operative in at least some contexts in which the content of the acceptance is practically relevant. In short, if my argument is correct, then everything I have said about having a rule suggests it is a certain kind of acceptance, and everything I have said about acceptance suggests the same.

Let me express the idea slightly more formally as the ‘acceptance theory’. According to the acceptance theory, that  $S$  has a rule that  $As$  ought to  $\phi$  means that enough members of  $S$  accept that  $As$  ought to  $\phi$  and their acceptances are belief-independent and practical. Similarly, that you have a personal rule that you are to  $\phi$  means that you accept that you ought to  $\phi$  and your acceptance is belief-independent and practical.

The most important virtue of the acceptance theory is that it fits with what we know about having a rule, but it has other advantages, too. I mentioned that the acceptance theory is a unified theory, unlike the protected reasons-based account of what it is to have a rule. It also preserves a connection with Hart’s theory of what it is to have a rule. There are different ways of understanding Hart’s theory, and I do not claim to have shown more than a similarity with his theory. Nonetheless, given the prominence of Hart’s theory in jurisprudence, the similarity is welcome. Unlike Hart’s theory, however, the acceptance theory is safe from

Warnock's Objection, which has been thought to cause so much trouble for Hart. Hart did not make clear the difference between his 'critical reflective attitude' and a belief in a reason to comply with a rule. The acceptance theory, in contrast, makes having a rule a matter of acceptance, and acceptance is distinguished from belief by its responsiveness to practical reasons, susceptibility to direct and voluntary control, and context-dependence.

The acceptance theory also connects having a rule, a traditional topic in jurisprudence, to presumptions and exclusionary rules, as understood in the law of evidence. I described, in Section 2, how the distinction between rules of thumb and strict rules is like the familiar distinction between rebuttable and irrebuttable presumptions. And in the next chapter I describe how having a rule is useful in some of the same ways as presumptions are useful. More generally, the acceptance theory builds a bridge between work in legal philosophy and work in other areas, including the philosophy of action. Indeed, perhaps the most significant implication of the acceptance theory is that having a rule is not a very distinctive phenomenon. If Bratman, Stalnaker, Cohen, and others are right, belief-independent acceptances are common. What distinguishes the acceptances in which having a rule consists is mainly their normative content. That is interesting because it suggests that we could learn more about rules by looking at situations in which similar acceptances play a role.

Having listed some of the theory's advantages, there are several points that need emphasizing.

*First point.* The acceptance theory only requires that 'enough' members of a society hold the appropriate acceptances for that society to have a rule. That is, it does not specify the proportion of members that must possess that attitude. Nor does it say for how long the members of that society must hold these acceptances. In fashioning the acceptance theory this way I have followed Hart, who said that these matters 'need not worry us more than the

question as to the number of hairs a man may have and still be bald'.<sup>23</sup> I should also be clear, again following Hart, that it is unnecessary that *every* member of a society have the right kind of acceptances for that society to have a rule. 'Enough' members means more than one, not everyone.

*Second point.* The acceptance theory is a theory of what it is to have *a* rule, not of what it is to have a strong rule or a weak rule, or a strict rule or a rule of thumb. However, it has the resources to explain what it is to have each type of rule. It is a matter of someone, or enough members of a society, accepting that the rule ought to be complied with in an all-things-considered or relative sense, or of their acceptances being operative or inoperative in certain contexts. For example, a society has a strong rule that *As* ought to  $\phi$  if and only if enough of its members accept that *As* ought to  $\phi$ , all things considered, and their acceptances are belief-independent and practical. Figure 4, in Section 2, sets out the associations between types of acceptances and types of rule-possession more fully.

The theory does not imply that every rule that a society has is strong or weak, or strict or a rule of thumb. Suppose that some members of a society accept that *As* ought to  $\phi$ , all things considered, while others accept that *As* ought to  $\phi$ , relative to some type of reason. Together, there are enough people who accept one or the other that the society as a whole has a rule that *As* ought to  $\phi$ . Nonetheless, there are not enough members of either kind for us to say that the society has a strong rule or a weak rule that *As* ought to  $\phi$ . The situation in Albania resembles this. In the mountainous north there was a strong rule against killing sheep

---

<sup>23</sup> Hart, *Concept of Law* (n 20) 56.

dogs; in the prosperous south there was a weak rule against doing so; in Albania there was simply *a* rule against killing sheep dogs.<sup>24</sup>

It is also possible for a society to have a strong rule and a weak rule, or a strict rule and a rule of thumb, with the same content. For instance, if half of the members of a society accept that *As* ought to  $\phi$ , all things considered, while the other half accepts that *As* ought to  $\phi$ , relative to some type of reason, this society has a strong rule and a weak rule that *As* ought to  $\phi$ . There is no contradiction here. It is not like saying this society has a rule that is both strong and weak, or that it both has and does not have a strong (or weak) rule. Rather, this society has two rules, distinguished by strength, but identical in content.<sup>25</sup>

*Third point.* The acceptance theory makes the possession of a rule turn exclusively on people's attitudes. It does not matter, in particular, whether people comply with the rule. This is one of the features of the acceptance theory that distinguishes it from Hart's theory. It is also what makes the acceptance theory well suited to account for the existence of some of the rules that I first mentioned in Chapter 2. The grypsmen, for instance, had a number of rules that applied to those who wished to join their caste, and which they themselves never complied with. That poses no problem for the acceptance theory, which asks only whether the grypsmen had the appropriate acceptances.

In another example from Chapter 2, a society has a rule in favour of shaking hands even after a plague arrives and most people begin to break that rule. The example causes problems for theories like Hart's, which insist on general compliance with a rule, but it can be explained according to the acceptance theory. People accept that they ought to shake hands,

---

<sup>24</sup> The same point could be made with respect to strict rules and rules of thumb. Interestingly, in either case, the analysis only applies to societal rules, not to personal rules. When it comes to personal rules, there is only one person's acceptance that matters. A personal rule, therefore, is either strong or weak, and either a strict rule or a rule of thumb.

<sup>25</sup> I am also assuming that 'enough' members of a society is no more than half, and that no members both accepts that *As* ought to  $\phi$  all things considered and accepts that *As* ought to  $\phi$  in a relative sense.

but in a relative sense. They accept that there is a reason to shake hands that is not overridden by reasons of convenience and the like, but that may be overridden by serious health considerations. Their acceptances lead people to shake hands before the plague arrives, but not after. Their level of compliance with the rule changes, but because their acceptances persist, so does the rule.

I should perhaps say, however, that the acceptance theory does predict that people will comply with their rules under certain conditions, namely, when they would do so if they believed that the rule ought to be complied with.

*Fourth point.* The acceptance theory says nothing about *why* anyone accepts or continues to accept that a rule ought to be complied with. To have a rule, it is enough that people *do* accept that. You might think that is too weak. Consider the fact that, while we tend to accept that forks ought to be placed to the left of the plate, few of us would do so if others did not as well. The same is true of many other social rules. It is likely true of rules that solve coordination problems, in particular.<sup>26</sup> Citing such examples, you might argue that, for a society to have a rule, it is not sufficient that a society's members accept that a rule ought to be complied with. Their acceptances must be held conditional on enough others holding the same acceptances.

One difficulty with this objection is that it is hard to think of supporting examples. That is, it is hard to think of a society that falls short of having some rule because its members' acceptances are not mutually dependent. Of course, that it is hard to think of such an example does not mean that none exist, but it does mitigate the scale of the problem, if indeed there is one. Also, supposing the objection is sound, the problem should be easy to fix.

---

<sup>26</sup> The classic work on rules (or norms) as solutions to coordination problems is David Lewis' *Convention: A Philosophical Study* (Harvard University Press 1969). See also: Edna Ullmann-Margalit, *The Emergence of Norms* (OUP 1977) ch 3; Eric Posner, *Law and Social Norms* (Harvard University Press 2000); Cristina Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge University Press 2006).

It would mean adding an extra condition setting out the way in which people's acceptances must be mutually dependent for them to have a rule. That would mean changing the acceptance theory, but the change would not be dramatic.

The more direct response is that my concern is the conditions under which a society has *a* rule, not when it has a rule of etiquette, or a rule that solves a coordination problem.<sup>27</sup> The acceptance theory is safe from the objection so long as it is *possible* for a society to have a rule without its members' acceptances being mutually dependent in some way. And it *does* seem to be possible. Take the taboo against cannibalism. If that taboo exists in a society, it is in virtue of the attitudes of its members. It seems possible that most members of such a society would not change their stance towards cannibalism just because others do. Most of them might continue to accept that cannibalism ought not to occur even if others no longer do. They would warn against the growing trend of cannibalism in society, criticize others' more lackadaisical response towards the threat, and urge a return to an earlier, less cannibal-friendly time. Maybe that is not how any society treats cannibalism, but it could be, and the possibility is enough to safeguard the acceptance theory.

## 9.5 CONCLUSION

Even though it suffers from well-recognized problems, Hart's thinking on the existence of social rules has dominated jurisprudential thinking on the topic for many years. Earlier in this thesis I added to the criticisms of Hart's theory. In this chapter I proposed an alternative

---

<sup>27</sup> David Lewis argues that at least some conventions (as solutions to coordination problems) are rules, but he identifies a number of other types of rules: Lewis (n 26) 97, 100-107. Edna Ullmann-Margalit argues for a theory of norms as solutions to certain types of social interaction problems, including coordination problems, but she clearly distinguishes the question 'why does a norm exist?' from the question 'under what conditions would we say a norm exists?' Her own approximate answer to the second question does not require that people who have a rule are disposed as they are because others are, too. See: Ullmann-Margalit (n 26) 8-13.

theory. The acceptance theory draws inspiration from Hart's work on social rules, but it avoids the problems his theory faces. It avoids them by making rule-possession turn solely on the kind of attitude Hart himself may have been seeking: an attitude different enough from belief to be responsive to practical reasons and capable of voluntary acquisition, but one that shares its motivational force. That attitude, I claim, is acceptance that the rule ought to be complied with, where the acceptance is practical and belief-independent.

# 10 Reasons, Possession, and Compliance

---

## 10.1 INTRODUCTION

In the last chapter, I offered an answer to the question, ‘what is it to have a rule?’, in the form of the acceptance theory. The theory says, essentially, that to have a rule is to accept that it ought to be complied with. The acceptances in which rule-possession consist are belief-independent, meaning they may not be accompanied by a belief with the same content. In this brief chapter, I try to explain how possessing this sort of acceptance may be justified, and thus why people bother having rules. I also make clear that the fact you ought to have a rule does not mean you ought to comply with it, on some or on any occasion.

## 10.2 REASONS TO HAVE RULES

### *Reasons to Act ‘As If’ and Reasons to Have Rules*

Why should someone have a rule? The question needs an answer because, as Warnock says,

there is always an alternative to having a rule, namely, leaving things, or leaving people, to go on as they may; and since it is difficult to see how having rules could be regarded simply in itself, for its own sake, as a good thing, there needs to be a reason for *not* just leaving things alone.<sup>1</sup>

---

<sup>1</sup> Geoffrey Warnock, *The Object of Morality* (Methuen & Co Ltd 1971) 36.

Warnock suggests that having a rule is rarely if ever of non-instrumental value. Whether or not that is true, it is not the usual reason for having a legal rule. The usual reason for having a legal rule is instrumental, and that is the sort of reason I will discuss.<sup>2</sup>

There is an instrumental reason to have a rule if having it will lead you to satisfy some other reason, one you would not have satisfied otherwise. Having a rule leads you to act as if you believe some action ought to be performed. You would have acted that way anyway if you actually believed the action ought to be performed. So, if having a rule is instrumentally justified, then you do not actually believe it ought to always be complied with, but having the rule leads you to act as if you did, which there is a reason to do. The acceptance theory says that, if you have a rule, you accept some action ought to be performed. Acceptance absent belief is what I termed ‘make-belief’. We can then say that an instrumental justification for having a rule is a justification for engaging in a kind of normative make-believe. You make-believe some action ought to be performed, without really believing that, in order to bring yourself to act as if you believe you did.

What distinguishes make-believe in the case of a rule from other make-believe is mainly its normative content. In Chapter 7, I gave examples of useful make-believe in other contexts, such as make-believing that God exists, that lions are not scary, and that a concentration camp is a game. In all these cases, you make-believe things are other than they are, to get by better in the world as it is. In Chapter 6, I gave examples of useful make-believe in evidence law. I showed how presumptions and fictions help to reduce deliberation costs and the costs of errors. In the rest of this section, I show how having a rule serves those same purposes, as well as helping to solve one other type of problem. To simplify matters, my examples are all of personal rules.

---

<sup>2</sup> I am only discussing reasons *for* having rules, not reasons *against* having rules. In each of the examples in this section I am claiming there is a reason to have a certain rule, but I leave it open whether you should all things considered have the rule.

*Jill's Case: Make-Belief and Efficiency*

Earlier I described what I termed 'ordinary decision-making'. It starts with you balancing all your relevant reasons, which leads you to form a judgment about what you ought to do, and from there to make a decision. Normally your decision brings you to do as you decided upon. There are some potential drawbacks to ordinary decision-making. In particular, going through the balancing process can be time-consuming, stressful, and the like. Having a rule, it is said, can help reduce some of the costs of deliberation.

Recall Jill's case from Chapter 5: Every year Jill goes on holiday and every year she has to again figure out where to go. Jill is aware of many pros and cons of each destination. To weigh them all would take a long time and involve a lot of hassle. But will Jill actually bother to go through the balancing process? Not necessarily. Suppose Jill believes she ought to always take her holidays in France, all things considered. In that case, she will not go through the balancing process. Why would she, if she thinks she already knows what it would tell her? From her perspective the weighing process is not only unpleasant but unnecessary. She will skip it entirely. Jill's belief that she ought to always go to France saves her from a costly process, making it a useful belief to have.

Suppose instead that Jill does *not* believe that she ought to always go to France for the holidays. Indeed, she does not believe that there is any one place where she ought to always take her holidays. In that case, Jill would save time and hassle by acting as she would were she to believe she ought to take her holidays in France. She has a reason to act as if she has that belief, even though actually she does not have that belief. She can bring herself to act that way by make-believing that she ought to take her holidays in France, or in other words, by having a strong rule that she ought to take her holidays there. If she did, she would treat the

matter as settled and deliberate no more about it, just as she would were she to believe the same. In this way make-believing that an action ought to be performed reduces the costs of deliberation.

Having a strong rule that she ought to go to France will certainly save Jill time and hassle, but it has its disadvantages. In particular, it will prevent Jill from responding when unusual or compelling reasons arise to go elsewhere. For example, if Jill's father were ill, her rule would prevent her from taking her holiday near him (unless, of course, he lived in France). The rule reduces Jill's deliberation costs, but by the same stroke, may lead her to make bad choices she could have avoided otherwise. It is possible (though we would have to know more to say for sure) that Jill would strike a better balance by having a weak rule that she ought to take her holidays in France. Perhaps she ought to just accept that she ought to go there relative to reasons of cost, pleasure, etc. That would leave her to deliberate about the weights of other types of reasons, if there are any. Having the weak rule would not reduce her deliberation costs as much as the strong rule would, but it would give her increased flexibility.

Whether Jill's rule is strong or weak, she also has the option of making it a rule of thumb. Given the nature of her goals, taking that option probably makes sense. For suppose Jill believes she ought to go somewhere else this year. Absent any rule, she would still skip the balancing process this year because she thinks she already knows what its result would be. There would be no advantage in terms of time saved or hassle avoided in her instead treating it as truth that she ought to go to France. I do not mean this is a conclusive reason for Jill to have a rule of thumb. There are other relevant considerations, such as how accurate Jill's judgment is. It does suggest, however, that for the purpose of reducing deliberation costs,

having a rule of thumb is just as good as having a strict rule, and this is in fact the common view.<sup>3</sup>

### *Black Widows: Make-Belief and Ignorance*

Jill's case demonstrates one of the potential problems with ordinary decision-making: completing the decision-making process may be costly. Another potential problem is that the process can be corrupted by incorrect or incomplete beliefs about the relevant reasons, leading to an incorrect judgment about what ought to be done, and from there to bad decisions and wrong actions. Having a rule can help avoid this problem with ordinary decision-making as well.

Recall the spider example from Chapter 5: Some spiders with red markings on their abdomens are harmless false black widows. Other spiders with the same red markings are real black widows, deadly to humans. The trouble is that false black widows and real black widows look much alike, even apart from their red markings. You are no expert about spiders, and sometimes their similarity confuses you. Sometimes you think a false black widow is a real black widow. Much more importantly, sometimes you think a real black widow is a false black widow. Your misidentifications will not necessarily put you in danger. It depends on what else you believe. Suppose you believe you ought to avoid all spiders with red markings. In that case your misidentifications will not matter much, in practical terms. Even if you misidentify a real black widow as a false black widow, you will still (correctly) judge that you ought to avoid it. Things get dangerous only when the belief that a spider is a false black

---

<sup>3</sup> See: Alan Goldman, *Practical Rules: When We Need Them and When We Don't* (Cambridge University Press, 2002) 34: 'When we apply a rule, we need not search for and investigate factors outside its scope. But for just this resource-saving function, rules of thumb ... can serve almost as well.'

widow would not, even in combination with your other beliefs, lead you to judge that you ought to avoid the spider. At that point you stand a real risk of suffering a deadly bite.

Given that you have one incorrect belief (that a real black widow is a false black widow) it would be better, practically speaking, to have another incorrect belief (that you ought to avoid all spiders with red markings) than not. It would be better because of how the combination of incorrect beliefs would dispose you to act (to avoid all spiders with red markings, including the real black widows). For that reason, even if you do *not* believe you ought to avoid all spiders with red markings, you have a reason to act as if you had that belief.<sup>4</sup> That gives you a reason to make-believe that you ought to avoid all spiders with red markings. Put another way, it gives you a reason to have a strong rule that you ought to avoid all spiders with red markings.

When discussing Jill's case, I said that Jill would probably be best off having a rule of thumb, because a rule of thumb would reduce the costs of her deliberations without restricting her ability to go somewhere other than France if she believes she ought to. Things are different when it comes to black widows. For suppose you do know quite a bit about spiders, such that if you are *certain* that a spider is a false black widow, it almost surely is a false black widow. But given the risks, even 'almost' is not good enough. Making the rule a rule of thumb would only be justified if at some degree of confidence your judgment about whether a spider is a false black widow was *always* accurate. That is only likely to be the case for true experts. For everyone else, the rule should not be a rule of thumb.

---

<sup>4</sup> See: Robert Stalnaker, *Inquiry* (MIT Press 1987) 93: 'When is it reasonable to accept something one believes is false? When one believes that it is *essentially* true, or close to the truth – as close as one can get, or as close as one needs to get for the purposes at hand. ... Sometimes these decisions are based on practical considerations.' Emphasis in original. Do you believe it is 'essentially true' that all spiders with red markings should be avoided? No. But you do believe that practical considerations mean it is close enough to the truth to be worth accepting.

*Tooth Brushing: Make-Belief and Small Gains*

Consider a final example, this time one of Philip Pettit and Geoffrey Brennan's. (For 'calculate' read 'weigh the pros and cons'.)

Suppose that I were to calculate after every meal whether to clean my teeth. On the negative side I would count the non-trivial cost of going to the trouble required. What would I count on the positive? The temptation is to say: the importance of the cleaning for my dental health. But the fact is that the significance of each individual cleaning for my oral well-being is negligible or next to negligible. No individual cleaning is sufficient to make the difference between having healthy teeth and not having healthy teeth; this is because dental health is a vaguely defined gestalt. Thus, other things being equal, calculation after every meal would always fail to elicit a walk to the bathroom; the result would be, bad teeth.<sup>5</sup>

It is not difficult to determine whether the balance of reasons favours teeth-brushing on a particular night: it does not. Nor is it time-consuming or labour-intensive to make the determination. The problem is rather that if I were to *always* decide how to act in the ordinary way I would never brush one's teeth and end up in poor dental health. To achieve dental health 'one must forswear calculation over individual actions. Better find a rule and stick to it. [Brush] after every meal, or [brush] every day, or [brush] when some regular chance even occurs. Do anything other than weigh the pro's and con's in every instance'.<sup>6</sup>

As Pettit and Brennan present the situation, I ought to brush my teeth on *all* occasions of some type (for example, after every meal) but not on *each* of those occasions (for example, not after this meal, or that meal, or ...). Assuming this is correct, the problem is how I can bring myself to brush my teeth regularly, if every time I deliberate about whether to brush them, I correctly judge that I should not. Given the discussion so far, my suggestion is predictable: If only I were to believe that I ought to brush my teeth after each meal, relative to reasons of convenience and dental health, there would be no problem. Since I do not have

---

<sup>5</sup> Philip Pettit and Geoffrey Brennan, 'Restrictive Consequentialism' (1986) 64 *Australasian Journal of Philosophy* 438, 452.

<sup>6</sup> Pettit and Brennan (n 5) 452.

that belief, I ought to make-believe that I ought to brush my teeth after each meal, relative to those reasons. That will lead me to treat it as true that I ought to brush my teeth after each meal.<sup>7</sup> The matter being closed, I will treat further deliberation as irrelevant. The make-belief will have a peremptory effect: it will cut off internal debate or argument about the proposition I make-believe is true.

Should my rule be a rule of thumb or not? I said that Jill ought to have a rule of thumb about where to go on holiday, and that an expert on spiders ought to have a rule of thumb about which spiders to avoid. But in this case a rule of thumb would be worthless. If Pettit and Brennan are right about the nature of teeth-cleaning, it is not only true that I ought to not brush my teeth on a particular occasion; it is obviously true. Supposing all I had was a rule of thumb, on virtually every occasion I would be sufficiently confident that I ought to not brush my teeth to set the rule aside. For the rule to have any force at all, it has to apply regardless of how confident I am that I ought to not now brush my teeth.

### *Summary*

This section showed that, on the acceptance theory, an instrumental reason to have a rule is a reason to make-believe some action ought to be performed. It brings you to act as if you had a belief which you lack, but which it would be useful to have. Make-believing is justified, in particular, as a means of reducing deliberation costs and the costs of errors, two of the traditional reasons for having rules.

---

<sup>7</sup> This is not too far off what Pettit and Brennan suggest. They advise cultivating a 'predisposition' to 'adhere to a maxim', such as cleaning your teeth after every meal. Pettit and Brennan (n 5) 443ff.

### 10.3 REASONS TO COMPLY WITH RULES

In the last section I showed that the acceptance theory partially answers the question, ‘why have a rule?’ Now I want to show that the theory also helps to answer the question, ‘why comply with a rule?’ The contribution is negative, in that the theory tells us how little whether you have a rule, or ought to have a rule, matters to whether you ought to comply with it.

Start by supposing that you have a rule and that it applies to you on some occasion. Should you comply with it? The acceptance theory says that you have a rule if you accept some action ought to be performed (and your acceptance is operative in the appropriate contexts, but we can leave that aside). You can accept a false proposition, including a false normative proposition. So, the fact you accept that an action ought to be performed does not entail that it ought to be performed. In other words, that you have a rule, and it now applies, does not entail that you ought to comply with it. Even those who reject the acceptance theory should answer the question ‘no’, because otherwise you could make it the case that you ought to do something just by adopting a rule that requires it.<sup>8</sup>

Suppose you have a rule and, what is more, you ought (all things considered) to have it as your rule. Does it now follow that you ought to comply with the rule, where possible? To consider the argument for answering ‘yes’ in its strongest form, assume that having any rule leads you to do exactly what it requires and nothing more. In that case, normally, if you ought to have a rule, then you ought to do as it requires. The reason is that, if a rule of yours requires you to do that which you ought not to do, there must have been a superior rule you could have had instead, like the first rule in every respect except that it does not require this action of you. The rule that is superior to every alternative rule, in this way, would not require

---

<sup>8</sup> There are well-known objections to bootstrapping reasons into existence like this. Those objections are most commonly discussed in relation to intention. The point was originally made by Michael Bratman in *Intentions, Plans, and Practical Reason* (first published 1987, CSLI Publications 1999) 24-27, 86-87. For a useful overview, see: Bruno Verbeek, ‘Rational Self-Commitment’ in Fabienne Peter and Hans Schmid, *Rationality and Commitment* (OUP 2007).

anything of you that you ought not to do, and would therefore be a rule you ought always to comply with.<sup>9</sup> (This argument only applies to rules justified on instrumental grounds, and even then cases like tooth-brushing are exceptions to it.)

For my purposes the more significant objection is that the starting assumption is often incorrect: often having a rule does *not* lead you to do as it requires and nothing more. There are several reasons why. First, you might not know whether a rule of yours applies. For example, to know whether the rule of etiquette against using a knife for a side dish applies, you first have to know whether this dish is a side dish, which is a complicated matter. Your ignorance creates a potential gap between what your rule requires and what you think it requires. Second, your rule might be a weak rule, designed to be overrideable. Indeed, the whole point of having a weak rule as opposed to a strong rule is that it will exert a more limited influence on your actions. Third, you might know your rule requires an action but by mistake fail to perform that action. Finally, just as akrasia disrupts the normal progression from a belief that ought to perform an action to its performance, an acceptance that you ought to do something does not guarantee you will do it.

The points in the last paragraph show there is a gap between what your rule requires and what it leads you to do, where only the latter is a source of instrumental reasons for having the rule. It is possible that you ought to have a rule even though it requires an action you ought not to perform, because on the whole it leads to better conformity with your reasons than any other rule would.

The black widow spider example provides a good illustration. In that example you have a rule, 'Avoid any spider with red markings'. On most occasions when the rule applies,

---

<sup>9</sup> I am assuming for simplicity's sake that you either ought or ought not to comply with a rule of yours. It may, however, be that there is neither a reason for nor against compliance with a rule. So, strictly speaking, the conclusion is that, if you ought to have a rule, then it is not the case that you ought not to comply with it, when compliance is possible.

you should not do as it says, because most spiders with red markings are harmless false black widows, not even worth the trouble of avoiding. Only on the rare occasions when a spider with red markings is a real black widow should you comply with the rule. The rule is highly overinclusive. But it has a great virtue: it is obvious when it applies. Red markings are easily spotted, so the odds are good that you will correctly judge when the rule requires something of you.

Consider an alternative to your simple rule:

Avoid any spider that (1) is black, (2) is two inches long and half an inch in diameter, and (3) has red markings in the shape of an hourglass on the underside of its abdomen.

This rule is not as overinclusive, but its relative accuracy comes at a price, because it is easier to make a mistake about what it requires. Is that spider black or just dark brown? Is it two inches long or more like three? Is that an hourglass or not? As the questions add up, so does the chance of answering one incorrectly. Other things being equal, the simpler rule is more likely to lead you to avoid a real black widow spider than the complicated rule is. Given the risks involved, this makes the simpler rule preferable, even though it is also more likely to lead you to avoid a false black widow.<sup>10</sup>

In the spider example, you ought to have the simple rule because you are able to accurately judge what it requires. In other cases, you ought to have a simple rule because it is less costly to figure out what it requires. Consider Jill's case again. She has to decide where to go on holiday every year. She can do that accurately by going through the balancing process, but going through the process is not worth it. Saving the time and hassle of weighing all the pros and cons of the various options more than makes up for being at a slightly worse destination. So Jill decides to adopt a rule. It is very simple, very blunt. It says: 'I ought to

---

<sup>10</sup> And the simpler rule is the preferable rule to have even if it leads you to do exactly as it requires.

always take my holidays in France'. It takes no time or effort to figure out what the rule requires, but it will often require her to go to France even though she ought to go elsewhere. For Jill, the trade is worth making, given how much she values efficiency in decision-making. Another rule might more closely reflect the balance of reasons on each occasion, but that accuracy will necessitate a more nuanced rule, and a more involved decision-making process.

#### 10.4 CONCLUSION

In this short chapter I considered two of the implications of the acceptance theory. The theory implies that instrumental reasons to have rules are typically reasons to make-believe that some action ought to be performed. You can be justified in make-believing some action ought to always be performed even though you would also be justified in not performing that action, on some occasion or on every occasion.

# Conclusion

---

Hart famously asked what it means for a mandatory rule to exist. Some people, he said, find the notion ‘utterly mysterious’. He aimed to dispel some of that mysteriousness through an account of what it means for a mandatory rule to exist in a society. In this thesis, I argued that Hart’s account is lacking, but that there is a better alternative available in the form of the acceptance theory. The acceptance theory answers the most puzzling part of Hart’s question, and so contributes to an important debate in jurisprudence.

My discussion of rules began with Hart’s practice theory. According to that theory, a society has a mandatory rule in virtue of a pattern of compliance with that rule among that society’s members and their possession of a certain attitude. I argued that this pattern of compliance is unnecessary. I gave as examples rules of a society that only apply to non-members, and rules that happen not to apply to any member of this society on any occasion. The lesson I drew was that a society has a rule simply in virtue of its members having the right attitude. The question, what makes it the case that a society has a rule?, therefore became, essentially, what attitude is the right attitude?

From Hart’s description of this attitude, others’ criticisms of Hart, and a discussion of partial conflicts between reasons, I concluded several things about the attitude on which having a rule depends. First, it leads you, as one of its possessors, to act as if you believe that rule ought to be complied with. Second, this attitude is not entailed by the belief that this rule ought to be complied with. Third, this attitude does not entail the belief that this rule ought to be complied with. There were other points, which I relied on to distinguish possession of

rules of various types, but these were the central ones. The task was thus to find an attitude with the motivational force of belief that something ought to be done, but that neither entails nor is entailed by that belief.

I looked for this attitude in the philosophy of action and the law of evidence. I described how philosophers of action have distinguished belief from another propositional attitude, namely, acceptance. Acceptance of a proposition leads its possessor to act in the same ways as a belief in that proposition normally would, but it neither entails nor is entailed by that belief. Scholars in the law of evidence use acceptance to help explain the nature of legal presumptions. The resources needed to account for rule-possession were already there, in legal doctrine and other areas of philosophy; they have just not been drawn on before.

With these resources in hand, I proposed the acceptance theory of rule-possession. I argued that the belief-independent acceptance that a rule ought to be complied with has the right motivational force, and the right relationships with the relevant normative beliefs. It is also under a person's voluntary control, and responsive to practical reasons, just as Hart seems to say is the attitude on which rule-possession depends. The acceptance theory builds on Hart's practice theory, but it does not suffer from the same problems as that theory does. It contributes to an answer to one of Hart's 'persistent questions' in jurisprudence, and it does so in part by drawing on work in other areas of law and philosophy.

The acceptance theory has the potential to improve our understanding of rules more generally. I argued that an instrumental justification for having a rule is a justification for engaging in a certain kind of normative make-believe. And I showed that, according to the acceptance theory, it is possible that you ought to break a rule you are nonetheless justified in possessing.

There are other ways of developing the acceptance theory, some of which I allude to in this thesis. For example, one of the unnecessary but common features of social rules is that they are sustained by interlocking acceptances. I gave rules of etiquette as an example. More important examples likely include rules of recognition and rules of customary law, including international customary law. It would be worthwhile to consider what extra conditions would be necessary to have such customary rules as opposed to a rule *simpliciter*. It would also be worth asking whether there are ‘shared’ acceptances, analogous perhaps to shared intentions or shared beliefs (such as there are).

The acceptance theory does not address permissive or power-conferring rules, or rules that are deliberately created. It provides a starting point for an analysis of them, however. Permissive rules, for example, are not so different from mandatory rules. With the success of the acceptance theory, it is reasonable to think that having a permissive rule is, or involves, accepting that an action may be performed. Some rules that are deliberately created (e.g., most legal rules) depend for their existence on social rules (e.g., the ultimate rule of recognition of a legal system). The acceptance theory could be the first step in an analysis of the existence conditions of such deliberately created rules. I also explained how the acceptance theory may indicate what a person must do to create a personal rule.

The acceptance theory can be extended in various directions, and it can be deepened. One way to deepen it would be to explore in more detail the connection between rules and acceptance-based notions including make-believe, pretense, imagination, and (non-legal) fictions. I explained how make-believe is related to the traditional reasons for having rules. To think of another obvious point of connection, systems of rules might be analogous to the games of imaginative pretense children play. The somewhat porous border between what children believe and make-believe could serve as a model of how our rules shape our beliefs, and how our beliefs shape our rule-based reasoning.

# Bibliography

- A Traveller, *The Art of Dueling* (Willoughby 1836) in Charles Haven, *Shooting Muzzle Loading Handguns* (Guns Inc 1947)
- Alexander L and Kress K, 'Against Legal Principles' (1997) 82 Iowa Law Review 739
- Alexy R, *A Theory of Constitutional Rights* (Julian Rivers tr, OUP 2002)
- Anonymous, *Modern Etiquette in Private and Public* (Frederick Warne 1872) 16-18 excerpted in Mary Douglas (ed), *Rules and Meanings: The Anthropology of Everyday Knowledge* (Routledge 1973)
- Audi R, 'Belief, Faith, and Acceptance' (2008) 63 International Journal of Philosophy of Religion 87
- Baier K, *The Moral Point of View: A Rational Basis for Ethics* (Cornell University Press 1958)
- Barber NW, *The Constitutional State* (OUP 2010)
- Bayles M, *HLA Hart's Legal Philosophy: An Examination* (Kluwer 1992)
- Bederman D, *Custom as a Source of Law* (Cambridge University Press 2010)
- Beltrán J, 'Legal Proof and Fact Finders' Beliefs' (2006) 12 Legal Theory 293
- Ben-Menachem H, 'Comment' in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy: The Influence of HLA Hart* (OUP 1987)
- Benditt T, *Law as Rule and Principle* (Stanford University Press, 1978)
- Bentham J, *A Treatise on Judicial Evidence* (JW Paget 1825)
- Bicchieri C, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge University Press 2006)
- 'The Fragility of Fairness: An Experimental Investigation of the Conditional Status of Pro-Social Norms' (2008) 18 Philosophical Issues 230
- Bix B, 'HLA Hart and the Hermeneutic Turn in Legal Theory' (1999) 52 SMU Law Review 167
- Black M, *Models and Metaphors* (Cornell University Press 1962)
- Bratman M, *Intentions, Plans, and Practical Reason* (first published 1987, CSLI Publications 1999)
- 'Practical Reasoning and Acceptance in a Context' (1992) 101 Mind 1
- Broome J, 'Normative Requirements' (1999) 12 Ratio 398
- 'Reasons' in R Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (OUP 2004)
- Buckareff A, 'Acceptance and Deciding to Believe', (2004) 29 Journal of Philosophical Research 173

Burros M, 'Plain and Simple; A Dish So Good Even the Fat is O.K.' *New York Times* (New York, 7 April 1993) < <http://www.nytimes.com/1993/04/07/garden/plain-and-simple-a-dish-so-good-even-the-fat-is-ok.html>> accessed 11 July 2011

Campbell K, 'Fuller on Legal Fictions' (1983) 2 *Law and Philosophy* 339

Canada Department of Fisheries and Forestry, *Canadian Fish Cook Book* (Canada Department of Fisheries and Forestry 1959)

Coady D, 'Preempting Preemption' in John Collins, Ned Hall, and LA Paul (eds), *Causation and Counterfactuals* (MIT Press 2004)

Cohen J, 'Belief and Acceptance', (1989) 98 *Mind* 367

-- *An Essay on Belief and Acceptance* (OUP 1995)

Coleman J, *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory* (OUP 2003)

Collins J, 'Preemptive Prevention' (2000) 97 *The Journal of Philosophy* 223

Collins J, Hall N, and Paul LA, 'Counterfactuals and Causation: History, Problems, and Prospects' in John Collins, Ned Hall, and LA Paul (eds), *Causation and Counterfactuals* (MIT Press 2004)

Dancy J, 'Enticing Reasons' in R Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (OUP 2004)

Davidson D, *Essays on Actions and Events* (OUP 1980)

Duff A, *Intention, Agency, and Criminal Liability* (Blackwell 1990)

Dundes AR and Dundes A, *Folk Law: Essays in the Theory and Practice of Lex Non Scripta* (University of Wisconsin Press 1995)

Durston G, *Evidence: Text and Materials* (2<sup>nd</sup> edn, OUP 2011)

Dworkin R, *Taking Rights Seriously* (Harvard University Press, 1978)

-- *A Matter of Principle* (Harvard University Press 1985)

Eggleston R, *Evidence, Proof, and Probability* (2<sup>nd</sup> edn, Wiedenfeld & Nicholson 1978)

Elster J, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge University Press 1983)

-- *The Cement of Society: A Study of Social Order* (Cambridge University Press 1989)

Endicott T, 'Are There Any Rules?' (2001) 5 *Journal of Ethics* 1999

Engel P, 'Believing, Holding True, and Accepting' (1998) 1 *Philosophical Explorations* 140

Engel P (ed), *Believing and Accepting* (Springer 2000)

Emson R, *Evidence* (5<sup>th</sup> edn, Palgrave Macmillan 2010)

Festinger L and Carlsmith J, 'Cognitive Consequences of Forced Compliance' (1959) 58 *The Journal of Abnormal and Social Psychology* 203

- Finlay S, 'What *Ought* Probably Means, and Why You Can't Deatch It' (2010) 177 *Synthese* 67
- Finnis J, 'Comment' in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy: The Influence of HLA Hart* (OUP 1987)
- Frankish K, *Mind and Supermind* (Cambridge University Press 2004)
- Franklin B, *The Autobiography of Benjamin Franklin* (J Bigelow ed, Putnam's 1916)
- Fuller L, 'Legal Fictions' (1930-1931) 25 *Illinois Law Review* 363, 513, 877 (published in three parts)
- Gardner J and Macklem T, 'Reasons' in Jules Coleman and Scott Shapiro (eds), *The Oxford Handbook of Jurisprudence & Philosophy of Law* (OUP 2004)
- Gendler T, 'On the Relation Between Pretense and Belief' in Domenic Lopes and Matthew Kieran (eds), *Imagination, Philosophy and the Arts* (Routledge 2003)
- Gibbard A, *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (OUP 1990)
- Gilbert M, 'Social Rules: Some Problems for Hart's Account, and an Alternative Proposal' (1999) 18 *Law and Philosophy* 141
- Goldman A, *Practical Rules: When We Need Them and When We Don't* (Cambridge University Press, 2002)
- Goldstein L, 'Moore's Paradox' in Pascal Engel (ed), *Believing and Accepting* (Springer 2000)
- Gottlieb G, *The Logic of Choice* (George Allen & Unwin 1968)
- Green L, 'Legal Positivism', *The Stanford Encyclopedia of Philosophy* (2003) <<http://plato.stanford.edu/entries/legal-positivism/>> accessed 6 July 2011
- Guttel E, 'Overcorrection' (2004) 93 *Georgia Law Journal* 241
- Hacker PMS, 'Hart's Philosophy of Law' in Hacker PMS and Raz J (eds), *Law, Liberty, and Morality: Essays in Honour of HLA Hart* (OUP 1977)
- Hart HM, Jr and Sacks AM, *The Legal Process: Basic Problems in the Making and Application of Law* (William N Eskridge, Jr and Philip P Frickey eds, Foundation Press, 1994)
- Hart HLA, 'Scandinavian Realism' (1959) 17 *Cambridge Law Journal* 233
- *Essays in Jurisprudence and Philosophy* (OUP 1983)
- *Essays on Bentham* (OUP 1982) 243
- *The Concept of Law* (2<sup>nd</sup> edn, OUP 1994)
- Hasluck M, *The Unwritten Law in Albania* (Cambridge University Press 1954)
- Heine B, 'The Mountain People: Some Notes on the Ik of North-Eastern Uganda' (1985) 55 *Africa* 3
- Heiner R, 'Imperfect Decisions and the Law: On the Evolution of Legal Precedent and Rules' (1986) 15 *Journal of Legal Studies* 227

- Ho H, *A Philosophy of Evidence Law* (OUP 2008)
- Honoré T, *Responsibility and Fault* (Hart 1999)
- Hurd H, *Moral Combat: The Dilemma of Legal Perspectivalism* (Cambridge University Press 1999)
- Huss B, ‘Three Challenges (and Three Replies) to the Ethics of Belief’ (2009) 168 *Synthese* 249
- James W, *The Varieties of Religious Experience* (first published 1902, Harvard University Press 1985)
- *The Principles of Psychology* (first published 1890, Cosimos 2007)
- Kaminski M, *Games Prisoners Play* (Princeton University Press 2004) 34
- Keane A, *The Modern Law of Evidence* (7<sup>th</sup> edn, OUP 2008)
- Kiesewetter B, ‘“Ought” and the Perspective of the Agent’ (2011) 5 *Journal of Ethics & Social Philosophy* 1.
- Kemp J, ‘Review of *The Concept of Law*’ (1963) 13 *Philosophical Review* 188
- Kenny A, ‘Practical Inference’ (1966) 26 *Analysis* 65
- Levinson S, ‘Who is a Jew(ish) Justice?’ (1989) 10 *Cardozo Law Review* 2359
- Lewis D, *Convention: A Philosophical Study* (Harvard University Press 1969)
- ‘Causation’ (1973) 70 *Journal of Philosophy* 556
- ‘Postscripts to “Causation”’ in David Lewis, *Philosophical Papers: Volume II* (OUP 1986)
- Lessig L, ‘The Regulation of Social Meaning’ (1995) 62 *University of Chicago Law Review* 943
- Lloyd D, *The Idea of Law* (Penguin Books 1964)
- Lyons D, *Ethics and the Rule of Law* (Cambridge University Press 1984)
- MacCormick N, *Legal Reasoning and Legal Theory* (OUP 1978)
- *HLA Hart* (2<sup>nd</sup> edn, Stanford University Press 2008) 56
- Macoun R, ‘Inside the Black Box: What Empirical Research Tells Us About Decision-Making by Civil Juries’ in Robert Litan (ed), *Verdict: Assessing the Civil Jury System* (Brookings Institution 1993).
- Marmor A, *Philosophy of Law* (Princeton University Press 2011)
- Mele A, *The Springs of Action* (OUP 1992)
- Melville H, *Typee: A Peep at Polynesian Life* (first published 1846, Northwestern University Press 1968)
- Mendonca D, ‘Presumptions’ (1998) 11 *Ratio Juris* 399

- Menzies P, 'Counterfactual Theories of Causation', *The Stanford Encyclopedia of Philosophy* (2008) <<http://plato.stanford.edu/entries/causation-counterfactual/>> accessed 8 July 2011
- Mill JS, *System of Logic, Ratiocinative and Inductive* (first published 1843, Kessinger Press 2004)
- Moore M, 'Authority, Law, and Razian Reasons' (1989) 62 *Southern California Law Review* 829
- Morgan E, 'Some Observations Concerning Presumptions' (1931) 44 *Harvard Law Review* 906
- Murphy P, *Murphy on Evidence* (11<sup>th</sup> edn, OUP 2009)
- Nietzel M, McCarthy D, and Kerr M, 'Juries: The Current State of the Empirical Literature' in Ronald Roesch, Stephen Hart, and James Ogloff (eds), *Psychology and Law: The State of the Discipline* (Kluwer 1999)
- Nozick R, *The Nature of Rationality* (Princeton University Press 1993)
- Olivier P, *Legal Fictions in Practice and Legal Science* (Rotterdam University Press 1975)
- Pascal B, *Pensées and Other Writings* (Honor Levi translator, OUP 1995)
- Patterson D, 'Explicating the Internal Point of View' (1999) *SMU Law Review* 67
- Payne M, 'Hart's Concept of a Legal System' (1976) 18 *William and Mary Law Review* 287
- Amanda Perreau-Saussine and James Bernard Murphy (eds), *The Nature of Customary Law* (Cambridge University Press 2007)
- Perry J, 'Belief and Acceptance' (1980) 5 *Midwest Studies in Philosophy* 533;
- Perry S, 'Hart on Social Rules and the Foundations of Law: Liberating the Internal Point of View' (2006) 75 *Fordham Law Review* 1171
- 'Holmes versus Hart: The Bad Man in Legal Theory' in Steven J Burton, ed, *The Path of Law and its Influence: The Legacy of Oliver Wendell Holmes, Jr* (Cambridge University Press 2000)
- Persson I, 'A Consequentialist Distinction between What we Ought to Do and Ought to Try' (2008) 20 *Utilitas* 348
- Pettit P and Brennan G, 'Restrictive Consequentialism' (1986) 64 *Australasian Journal of Philosophy* 438
- Pettit P, 'Hope and its Place in the Mind' (2004) 592 *Annals of the American Academy of Political and Social Science* 152
- Posner E, *Law and Social Norms* (Harvard University Press 2000)
- Post E, *Etiquette in Society, in Business, in Politics and at Home* (Funk and Wagnalls Company 1922)
- Postema G, 'Coordination at the Foundations of Law' (1982) 11 *The Journal of Legal Studies* 165
- Raz J, 'Legal Principles and the Limits of Law' (1972) 81 *Yale Law Journal* 823

- 'Reasons, Requirements, and Practical Conflicts' in Stephen Korner (ed), *Practical Reason* (Yale University Press 1974) 26. Raz summarizes his discussion of partial conflicts in *Practical Reason and Norms* (n 5) 202 n 4.
- 'Reasons for Action, Decision, and Norms' (1975) 84 *Mind* 481
- 'Promises and Obligations' in Peter Hacker and Joseph Raz (eds), *Law, Morality and Society: Essays in Honour of HLA Hart* (OUP 1977)
- 'Introduction' in Joseph Raz (ed), *Practical Reasoning* (OUP 1978)
- *The Authority of Law: Essays on Law and Morality* (OUP 1979)
- *The Concept of a Legal System* (2<sup>nd</sup> edn, OUP 1980)
- 'Facing Up: A Reply' (1989) 62 *Southern California Law Review* 1153
- *Practical Reason and Norms* (2<sup>nd</sup> edn, Princeton University Press 1990)
- *Engaging Reason: On the Theory of Value and Action* (OUP 1999)
- 'Personal Practical Conflicts', in Peter Bauman and Monika Betzler (eds), *Practical Conflicts: New Philosophical Essays* (Cambridge University Press 2004),
- 'Reasons: Explanatory and Normative' in Constantine Sandis (ed), *New Essays on the Explanation of Action* (Palgrave Macmillan, 2009)
- 'Reasons: Practical and Adaptive' in David Sobel and Stephen Wall (eds), *Reasons for Action* (Cambridge University Press 2009)
- 'On the Guise of the Good' in Sergio Tenenbaum (ed), *Desire, Practical Reason, and the Good* (OUP 2010)
- Rescher N, *Presumption and the Practices of Tentative Cognition* (Cambridge University Press 2006)
- Ross A, 'Review of *The Concept of Law*' (1962) 71 *Yale Law Journal* 1185
- Sainsbury M, 'Fiction, and Acceptance-Relative Truth, Belief, and Assertion' in Franck Lihoreau (ed), *Truth in Fiction* (Transaction Books 2011)
- Sartorius R, 'Positivism and the Foundations of Legal Authority' in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy: The Influence of HLA Hart* (OUP 1987)
- Setiya K, *Reasons without Rationalism* (Princeton University Press 2007)
- Schaffer J, 'Trumping Preemption' (2000) 97 *The Journal of Philosophy* 165
- Schauer F, 'Exceptions' (1991) 58 *University of Chicago Law Review* 871
- *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (OUP 1991)
- 'Prescriptions in Three Dimensions' (1997) 82 *Iowa Law Review* 911
- 'On the Supposed Defeasibility of Legal Rules' (1998) 51 *Current Legal Problems* 223

- Schwartz W, Baxter K, and Ryan D, 'The Duel: Can These Gentlemen Be Acting Efficiently?' (1984) 13 *Journal of Legal Studies* 321
- Shapiro S, 'On Hart's Way Out' (1998) 4 *Legal Theory* 469
- and McClennen E, 'Rule-Guided Behaviour' in Peter Newman (ed), *The New Palgrave Dictionary of Economics and the Law* (Palgrave Macmillan 1998)
- 'The Bad Man and the Internal Point of View', both in Steven J Burton, ed, *The Path of Law and its Influence: The Legacy of Oliver Wendell Holmes, Jr* (Cambridge University Press 2000)
- 'What is the Internal Point of View?' (2006) 75 *Fordham Law Review* 1157
- 'How Rules Affect Practical Reasoning' in Bruno Verbeek (ed), *Reasons and Intentions* (Ashgate 2008)
- Sherwin R, 'Opening Hart's concept of Law' (1986) 20 *Valparaiso University Law Review* 385
- Shwayder D, *The Stratification of Behaviour* (Brill 1971)
- Smullyan R, *This Book Needs No Title: A Budget of Living Paradoxes* (Prentice Hall 1980)
- Soper P, 'Law's Normative Claims' in Robert P George (ed), *The Autonomy of Law* (OUP 1996)
- Stalnaker R, *Inquiry* (MIT Press 1987)
- 'Common Ground' (2002) 25 *Linguistics and Philosophy* 701
- 'On "Moore's Paradox"' in Pascal Engel (ed), *Believing and Accepting* (Springer 2000)
- Stendhal (Beyle H-M), *The Red and the Black* (first published 1830, translated by Burton Raffel, Modern Library 2003)
- Stephen JF, *A Digest of the Law of Evidence* (12<sup>th</sup> edn, Macmillan 1946)
- Stocker M, 'The Schizophrenia of Modern Ethical Theories' (1976) 73 *Journal of Philosophy* 453
- Streumer B, 'Can Consequentialism Cover Everything?' (2003) 15 *Utilitas* 237
- Sunstein C, 'Political Conflict and Legal Agreement' (The Tanner Lectures on Human Values, delivered at Harvard University 1994)  
<<http://www.tannerlectures.utah.edu/lectures/documents/Sunstein96.pdf>> accessed 7 July 2011
- Tapper C, 'A Note on Principles' (1971) 34 *Modern Law Review* 628
- Taylor J, *Hoppin' John's Low Country Cooking: Recipes and Ruminations from Charleston and the Caroline Coastal Plain* (Houghton Mifflin Harcourt 2000)
- Thayer JB, *A Preliminary Treatise of Evidence at the Common Law* (Little, Brown, & Co 1898)
- Thomson JJ, *Goodness and Advice* (Amy Gutman ed, Princeton University Press 2001)
- Toh K, 'Hart's Expressivism and his Benthamite Project' (2005) 11 *Legal Theory* 75

- 'Raz on Detachment, Acceptance and Describability' (2007) 27 *Oxford Journal of Legal Studies* 403
- 'An Argument Against the Social Fact Thesis (And some Additional Preliminary Steps Towards a New Conception of Legal Positivism)' (2008) 27 *Law and Philosophy* 445
- Turnbull C, *The Mountain People* (Jonathan Cape 1972)
- Twining W and Miers D, *How to do Things with Rules* (5<sup>th</sup> edn, Cambridge University Press 2010)
- US Courts for the Sixth Circuit, *Pattern Criminal Jury Instructions* (2009) issued by the U.S. Courts for the U.S. Sixth Circuit <[http://www.ca6.uscourts.gov/internet/crim\\_jury\\_insts.htm](http://www.ca6.uscourts.gov/internet/crim_jury_insts.htm)> accessed 10 July 2011
- U.S. Courts for the Ninth Circuit, *Manual of Civil Model Jury Directions* (2007) <<http://www.lb9.uscourts.gov/>> accessed 10 July 2011
- Ullmann-Margalit E, *The Emergence of Norms* (OUP 1977)
- 'On Presumption' (1983) 80 *The Journal of Philosophy* 143
- and Margalit A, 'Holding True and Holding as True' (1992) 92 *Synthese* 167
- Unwin N, 'What Does it Mean to Aim at Truth?' (2003) 40 *American Philosophical Quarterly* 91
- van Fraassen B, *The Scientific Image* (OUP 1980)
- Velleman D, 'On the Aim of Belief' in *The Possibility of Practical Reason* (OUP 2000)
- Verbeek B, 'Rational Self-Commitment' in Fabienne Peter and Hans Schmid, *Rationality and Commitment* (OUP 2007)
- von Wright GH, *Norm and Action* (Routledge & Kegan Paul 1963)
- Walton K, *Mimesis and Make-Believe* (Harvard University Press 1990)
- 'Metaphor and Prop Oriented Make-Believe' (1993) 1 *European Journal of Philosophy* 39
- Warnock G, *The Object of Morality* (Methuen & Co Ltd 1971)
- Wells H, 'The End of the Affair: Anti-Dueling Laws and Social Norms in Antebellum America' (2001) 54 *Vanderbilt Law Review* 1805
- Wigmore JH, *A Treatise on the System of Evidence at Common Law: Including the Statutes and Decisions of all Jurisdictions of the United States* (vol 4, Little, Brown, & Co. 1905)
- Williams B, 'Deciding to Believe' in *Problems of the Self* (Cambridge University Press 1976)
- *Truth and Truthfulness: An Essay in Genealogy* (Princeton University Press 2002)
- Williams J, *Dueling in the Old South: Vignettes of Social History* (Texas A&M University Press 2000)
- Wilson G, *Intentionality of Human Action* (Stanford University Press 1989)

Wozley AD, 'The Existence of Rules' (1967) 1 Nous 63

Zimbardo P, *The Lucifer Effect* (Random House 2008)