11 Enforcing Regulation

Applying regulatory controls on the ground involves the carrying out of a number of tasks. Even if it is assumed that regulatory objectives have been established with clarity, there is still much work to do. Information on risks, harms, and mischief creators has to be gathered, rules and other responses have to be devised, and the behaviour of regulatees has to be modified. An astute regulator, moreover, will also set about assessing whether current efforts are achieving the right results and whether there is a need to change strategy in order to improve performance. It is accordingly useful to think of enforcement, in its narrow 'behaviour modification' sense, as just one part of the broader regulatory process—a process that involves five core tasks.[1] These tasks can be set out as in the DREAM framework of Figure 11.1.

This chapter uses the DREAM framework to organize a discussion of the main challenges that regulators encounter in seeking to apply enforcement on the ground.

1.	**DETECTING**	The gaining of information on undesirable and non-compliant behaviour.
2.	**RESPONDING**	The developing of policies, rules, and tools to deal with the problems discovered.
3.	**ENFORCING**	The application of policies, rules, and tools on the ground.
4.	**ASSESSING**	The measuring of success or failure in enforcement activities.
5.	**MODIFYING**	Adjusting tools and strategies in order to improve compliance and address problematic behaviour.

Figure 11.1. Regulatory tasks: the DREAM framework

[1] See R. Baldwin and J. Black, *Defra: A Review of Enforcement Measures and an Assessment of their Effectiveness in Terms of Risk and Outcome* (London, 2005).

Detecting: Identifying Non-compliant and Undesirable Behaviour

Uncovering undesirable behaviour through detection is a first step in regulatory enforcement. Detection challenges are, however, often severe. Enforcers frequently face extreme difficulties in detecting errant behaviour when the regulated community is extensive (as where certain environmental controls cover the whole population) and where breaching rules is cheap and easily carried out in a clandestine manner. Resourcing realities often mean that enforcers have to rely on tip-offs from the public or hotlines and whistleblowing processes. As a result, the regulators will receive a good deal of unreliable information and will commonly be ill-placed to calculate the real level of 'off-the-screen' activity that detracts from the achievement of objectives.

In some areas it is, therefore, extremely difficult to state what 'compliance' involves and the problem of constructing an agreed understanding can be bedevilled by legal uncertainties. The latter sometimes stem from drafting weakness, but divergences of understanding between the judiciary and the regulators can also prove a problem—notably regarding the purposes and objectives of the regulation at issue. In cases where there are unresolved disagreements on the meaning of compliance, this renders detection activity extremely fraught.

Resourcing constitutes a perennial constraint on detection. In many controlled areas the calculation of levels of compliance and the incidence of 'off-the-screen' activity would demand the operation of registration schemes or the carrying out of surveys, but funds may not permit such activities. In other domains, such surveys are conducted and, in many sectors, programmes of random inspection are used to obtain relevant data. In yet other areas, detection can only be carried out after the event, and this impedes precautionary intervention.

In responding to these challenges, regulators must first develop clear conceptions of their aims and an appropriate disaggregation of those objectives into subsidiary aims so that achievable targets can be set and problems identified in a manageable way.[2] If this is not done, the regulators will not know what sort of errant behaviour they need to detect. Regulatory objectives, moreover, may change over time and, in addition, the threats to achieving objectives may shift continuously. A regulator of, say, fisheries will thus have to deal with changes in priorities regarding the protection of different species of fish (or regarding protecting fish versus protecting employment); it will also face emergent risks from innovative fishing technologies and new fishing enterprises.

[2] M. Sparrow, *The Regulatory Craft* (Washington, DC, 2000), 146–9.

Second, in such a state of flux it is essential to be able to identify levels and patterns of compliance, but change poses challenges. New methods of avoiding the rules or concealing non-compliance may be devised constantly. Enterprises may be creatively complying with, or breaking, the rules in innovative ways. A given set of rules or a licensing regime may be impacting on enterprises less than it did formerly. A regulator, accordingly, needs to be able to detect not merely the levels of any non-compliance with requirements but also the extent of any 'off-the-screen' or 'invisible' black market activity that affects the achieving of the agency's legitimate objectives. Third, regulators have to assess the extent to which compliance with the relevant legal requirements will not be enough to achieve agency objectives. In a world of change, with new problems and strategies for escaping the rules, it is essential to know, in a continuing manner, the gap between rules and objectives. How these challenges can be met will be returned to in the next chapter.

Responding: Developing Rules and Tools

A second core task of regulatory enforcement is the development of those rules and tools that are fit for purpose both in detecting non-compliance or undesirable behaviour and in producing compliance with relevant requirements. Although the potential list of enforcement tools is long,[3] not every regulator has the full complement, or indeed anything approaching it. This is no peripheral matter. The absence of a relevant tool—such as a power to take samples or to fine on the spot—may be seen by enforcers as a significant impediment to effective control.[4]

Enforcers are, however, often constrained in their development of new tools by a number of factors—including institutional environments. Legislation may often be needed in order to introduce new powers, and it is common for officials to consider that new legislation (even secondary legislation) is an unrealistic political prospect. Existing bodies of legislation (particularly European Directives) are often seen as constraints on the use of new tools and uncertainties in legislative requirements tend to blight creative approaches to new tools. Government policies and institutional factors are also often seen as

[3] In Defra alone over 40 different powers are deployed. Main types of enforcement tool are: tools relating to the *continuation of business/operations* (e.g. licence revocation powers); *monetary or financial tools* (e.g. fines); *restorative tools* (remediation orders; restorative conferences); *undertaking and compliance management tools* (e.g. voluntary or enforceable undertakings); *performance disclosure instruments* (e.g. league tables). These are in addition to pre-enforcement tools such as warnings, notices, etc.—see Baldwin and Black, *Defra: A Review of Enforcement Measures.*

[4] See, e.g., Defra, *Fly-Tipping Strategy* (London, 2004).

an impediment to new tool use—especially when these involve dispersions of regulatory responsibility across numbers of bodies or where attention is directed at enforcing existing tools to the detriment of forward looks at new powers. Resource constraints, as ever, constitute a hurdle—especially where these stand in the way of the surveying or inspection exercises that are needed in order to reveal the true incidence of non-compliance or unwanted activity and hence the need for new tools and strategies.

RULES AND ENFORCEMENT

In developing regulatory rules, it should be borne in mind that the type of rule adopted may impact on enforceability and the achieving of objectives. Not all kinds of rule can be enforced with the same degree of success.[5] Rules may fail for a number of reasons, for example, because they are too vague or too long and complex to understand readily or to enforce; or because they prohibit desirable behaviour or they do not cover certain undesirable conduct. Different regulatory contexts, furthermore, may demand rules with different qualities or dimensions.

Rules may vary according, *inter alia*, to: degree of specificity or precision; extent, coverage, or inclusiveness; accessibility and intelligibility; legal status and force; and the prescriptions or sanctions they incorporate.[6] Rules, moreover, have to be employed by enforcers in conjunction with different compliance-seeking strategies—be these prosecutions, administrative sanctions, or processes of persuasion, negotiation, advice, education, or promotion.

Different enforcement strategies may thus call for different kinds of rule. If prosecutions are the main mode of enforcement, this may call for precise rules so that guilt or innocence can be established easily. (As a result, these rules may be long and complex.) If broad promotion of good practice is to be used (e.g. in leaflets or guidance), then less precise but more accessible rules may be more effective.

As to the selection of enforcement strategies (and, accordingly, accompanying rule types) it has been argued that this requires an analysis of the kinds of regulatee being dealt with.[7] If the regulatee is well-intentioned (i.e. wishes to comply) and is ill-informed (about legal requirements or how to meet these),

[5] See generally, R. Baldwin, *Rules and Government* (Oxford, 1995); id., 'Why Rules Don't Work' (1990) 53 *MLR* 321; J. Black, *Rules and Regulators* (Oxford, 1997); id., 'Using Rules Effectively', in C. McCrudden (ed.), *Regulation and Deregulation* (Oxford, 1999); C. Hood, *Administrative Analysis: An Introduction to Rules, Enforcement and Organisations* (Brighton, 1986); C.S. Diver, 'The Optimal Precision of Administrative Rules' (1983) 93 *Yale LJ* 65.

[6] For other approaches to the dimensions of rules, see Black, *Rules and Regulators*, 21 and Diver, 'Optimal Precision'.

[7] See Baldwin, *Rules and Government*, ch. 6; I. Ayres and J. Braithwaite, *Responsive Regulation* (Oxford, 1992), ch. 2.

prosecution may be a lower priority than educating and promoting—since information rather than a big stick is required. Accessible rules will, accordingly, be useful. The well-intentioned, well-informed regulatee will be able to cope with more detailed rules. The ill-intentioned, ill-informed category may demand a higher level of prosecution and, accordingly, precise rules will be in order. Finally, the ill-intentioned, well-informed regulatee will demand strategies, rules, and sanctions that can cope with deliberate rule avoidance, and mixtures of general and specific rules may be appropriate.[8]

Effective enforcement thus calls for judgements to be made concerning blends of enforcement strategies and the rule types that will best produce compliance. This suggests that informers and rule-makers should ask the following questions:[9]

- What is the undesirable behaviour, or mischief at issue?
- Who is responsible for the mischief?
- Which enforcement strategies will best lead the mischief creators to comply?
- Which types of rules best complement those strategies?

Such an approach presupposes that types of regulatee can be identified in the various sectors regulated. This will allow strategies and rules to be designed accordingly. If this is not possible or sectors have numbers of different kinds of regulatee, it may be necessary for agencies to equip their enforcers with an array of rules and strategies to cope with all eventualities. This is very costly in rule-making resources. It can similarly be cautioned that pyramidic approaches to enforcement involve progressing through various strategies in a serial fashion and this may also make large demands on rule-making resources, since those different strategies should ideally be matched with different types of rule. When types of regulatee can be identified, a specific and a targeted, rather than a pyramidic, enforcement strategy may constitute a more effective use of resources.[10]

[8] See the discussion of creative compliance below and R. Kagan and J. Scholz, 'The Criminology of the Corporation and Regulatory Enforcement Strategies', in K. Hawkins and J. Thomas (eds), *Enforcing Regulation* (Boston, 1984)—who distinguish between corporations who are 'amoral calculators', 'political citizens', and 'organizationally incompetent'.

[9] See Baldwin, *Rules and Government*, ch. 6; id., 'Why Rules Don't Work'; and id., 'Governing with Rules', in G. Richardson and H. Genn (eds), *Administrative Law and Government Action* (Oxford, 1994) for a discussion of these questions and reasons why rule-makers may fail to adopt such an approach to rule-making. (They tend to assume enforcement is unproblematic and do not seek information on enforcement; they tend to underestimate the political problems involved in making rules; and they are subject to disruptive political pressures from within and beyond the organization.)

[10] See Baldwin, *Rules and Government*, 158 n. 25 and Ayres and Braithwaite, *Responsive Regulation*, ch. 2. On pyramidic approaches, See Chapter 12 below.

CREATIVE COMPLIANCE

Even if compliance with rules is achieved by enforcers, this is not the end of the story. Desired objectives may not be achieved for two main further reasons. The first of these is what has been termed 'creative compliance'.[11] This is the process whereby those regulated avoid having to break the rules and do so by circumventing the scope of a rule while still breaching the spirit of the rule. Let us suppose that, in order to protect small shops, a government legislates to prohibit retail premises with over 10,000 square metres of floor space from opening on Sunday afternoons. A retail firm might creatively comply with such a rule by dividing its 12,000 square metre operation into two linked operations of 6,000 square metres. It complies with the law but avoids the thrust of the legislation.

In some fields (e.g. taxation), whole industries are devoted to creative compliance and the challenge for regulatory rule-makers and enforcers is to devise ways to keep the problem under control. This may be difficult for a number of reasons. As McBarnet and Whelan note, regulated industries may apply political pressure to regulators and demand detailed rules so that the rule of law and principles of certainty are served, but such types of rules may in reality be the very formulations that are most easily side-stepped by creative compliers. One response is to reinforce detailed rules with open-textured and general rules that are more difficult to circumvent.[12]

INCLUSIVENESS

A second reason why successfully achieving compliance may still fail to produce desired results is that ill-formulated rules may prove over- or under-inclusive.[13] They may discourage desirable behaviour or fail to prevent undesirable behaviour. Bardach and Kagan[14] suggest that regulators tend to over-regulate with over-inclusive rules for a number of reasons. Amongst these are, first, the costs of gaining the information necessary for targeting rules perfectly. These costs can be very high, and rule-makers tend to solve the problem by writing over-inclusive rules and relying on selective enforcement. (This conveniently shifts costs from rule-makers to enforcers.) Second,

[11] See D. McBarnet and C. Whelan, 'The Elusive Spirit of the Law: Formalism and the Struggle for Legal Control' (1991) 54 *MLR* 848; id., 'Challenging the Regulators: Strategies for Resisting Control', in C. McCrudden (ed.), *Regulation and Deregulation*; D. McBarnet, 'Law, Policy and Legal Avoidance' (1988) *Journal of Law and Society* 113.

[12] See pp. 305–6 below, and, for example, the general duties for employers set out in sections 2–9 of the Health and Safety at Work Act 1984, which may catch employers who creatively comply around more precisely formulated regulations on workplace health and safety.

[13] See Black, *Rules and Regulators*, 7–10.

[14] E. Bardach and R. Kagan, *Going by the Book* (Philadelphia, 1982), 66–77.

rule-makers tend to opt for broad-brush solutions to problems. Third, pressure to avoid regulatory discretions and produce equal treatment under law tends to trade off efficiency for more rules. Finally, working on a problem while the political iron is hot tends to be necessary for rule-makers and the resultant tight deadlines rule out the precise targeting of rules.

Under-inclusiveness, on the other hand, may also result from informational problems. Thus, a rule may fail to come to grips with certain hazards because the regulator has not been able to develop the information necessary to identify the cause of the hazard.

Moreover, rules that deal with problems of inclusiveness, or coverage, may give rise to other problems. Colin Diver explains this well by supposing the need for a rule to stop airline pilots from flying when the social cost of allowing them to continue flying exceeds the social benefits of not having to replace them.[15] He suggests three formulations for such a rule.

Model I No person may pilot a commercial aircraft after their sixtieth birthday.

Model II No person may pilot a commercial aircraft if they pose an unreasonable risk of an accident.

Model III No person may pilot a commercial aircraft if they fall within one of the following categories. (There follow tables giving combinations of values for numerous variables including years, levels of experience, age, height, weight, blood pressure, heart-rate, eyesight, and a host of other medical factors affecting pilot performance.)

Model I is the most transparent and accessible rule. It is easily understood and easy to enforce, but gives rise to problems of inclusivity. Some pilots aged over 60 may present lower risks to passengers than some (unhealthy) pilots aged under 60. Model II offers a response to inclusivity (it states the rule's purpose) but, though accessible on its face, is vague, lacks clarity, and, for this reason, is difficult to enforce because it needs to be fleshed out and made precise. This is liable to prove an expensive and legally fraught process. Model III scores well on inclusivity and it precisely identifies hazard-causing factors (provided that it is constantly revised and supported by research on health risks). It is likely, however, to be very lengthy, technically complex, and difficult to apply without expert training or the hiring of specialized consultants.[16] It will also be extremely expensive for rule-makers to write, to keep up to date, and to locate within the necessary agreements on an exhaustive list of hazard-creating medical conditions. Model III is also likely to give rise to greater problems of creative compliance than Models I and II.[17]

[15] See Diver, 'Optimal Precision'.

[16] Model III may also give rise to issues of rule-choice. Which is the relevant rule to apply may become an issue in complex regimes.

[17] Thus a pilot suffering from a medical condition on the prohibited list might take a drug to remove that condition, but the drug might create another dangerous—but unlisted—condition.

Problems of over- and under-inclusiveness can also be approached by considering when it is better, in the face of uncertainty, to err on the side of over- or under-inclusiveness in regulating. Shrader-Frechette examines this issue in asking whether it is better to err by prohibiting the use of a technology that is falsely seen as dangerous but is really acceptable and safe (a 'Type I' error) or by allowing the use of a technology that is falsely seen as safe but which is really unsafe (a 'Type II' error).[18]

Type I errors are sometimes referred to as 'producer risks'[19] and Type II as 'consumer risks'. In practice, Shrader-Frechette contends, risk assessors tend to err on the side of avoiding Type I errors for a number of reasons: because pure science researchers prefer to suppose that no connection exists than to posit an effect (e.g. that a substance causes cancer);[20] because producers are seen as enjoying something analogous to a 'presumption of innocence' that places the burden on those asserting a harmful effect; and because many risk assessments are conducted by persons closely associated with, and sympathetic to, the product or technology at issue.[21]

Shrader-Frechette's argument, however, is that it is better, when uncertain, to err on the side of avoiding Type II errors—better to protect consumers rather than defend the producer's rights to sell products—because the burden of proof regarding risk acceptability should be placed on the person wishing to reduce producer rather than consumer risks.[22] She puts forward a number of reasons for this suggested approach:

[18] See K.S. Shrader-Frechette, *Risk and Rationality* (Berkeley, 1991). Such issues arise in making decisions on whether to regulate an area at all, as well as when deciding whether to draft an anticipated rule to cover a particular activity.

[19] Ibid., 132.

[20] See C.F. Cranor, *Regulating Toxic Substances* (New York, 1993), ch. 4. It has been argued elsewhere that less rigorous standards of proof are typically required to prevent the possibility of a Type II error than a Type I error—see G. Brennan, 'Civil Disaster Management: An Economist's View' (1991) 64 *Canberra Bulletin of Public Administration* 30–3. Stephen Tindale argues that the precautionary principle ('giving the environment the benefit of any reasonable doubt') has been seen in operation in some areas (e.g. new medicines or substances liable to enter the human food chain such as drinking water or meat containing growth hormones), but the Panglossian principle (optimism in the face of worrying evidence and the placing of the burden on those seeking to demonstrate that a risk arises) has also been encountered, notably in the environmental area (e.g. global warming, dog faeces, leukemia clusters around Sellafield, pesticides, and lead pollution). The prevalence of the Panglossian approach, Tindale says, leads to the undermining of respect for politics and those in authority—see S. Tindale, 'Procrastination and the Global Gamble', in J. Franklin (ed.), *The Politics of Risk Society* (Cambridge, 1998), and also A. Jordan and T. O'Riordan, 'The Precautionary Principle in UK Environmental Law and Policy', in T. Gray (ed.), *UK Environmental Policy in the 1990s* (Basingstoke, 1995).

[21] Beck argues that in 'risk society' in general, technological advances create new risks far more rapidly than conventional democratic mechanisms can devise responses—this would imply a tendency to under-regulate and to regulate in an under-inclusive manner (see U. Beck, 'The Politics of Risk Society', in Franklin, *Politics of Risk Society*).

[22] Beck also argues that the burden of proof should rest on risk creators to prove safety—'The Politics of Risk Society', 21.

- It is more important to protect from harm than to enhance welfare.
- Producers reap most of the benefits of a new technology; they should accordingly bear most of the risks and costs.
- Consumers merit greater protection than industry, since they have less information and fewer resources with which to deal with hazards.
- Lay persons should be accorded legal rights to bodily security—to minimize industry risk on efficiency grounds offends notions of such security and would be morally offensive.
- Producers may not always be able to compensate persons harmed by their products; it is better, accordingly, to err on the side of eliminating harms at source.
- If there is uncertainty about the level of harm, it is difficult to justify imposing a risk on consumers.
- On democratic grounds there ought to be no imposition of risks without the informed consent of those who are to bear the risks.
- If consumers have not given informed consent, industry ought to bear the burden of proving that imposing a consumer risk is justified.
- Minimizing consumer risk is less likely to threaten social and political stability than minimizing producer risk.

Contrary to the above approach, it can be contended that even if one accepts the value of consumer protection, there are a number of reasons why one might, on occasions, want to favour producer rather than consumer protection and accordingly err on the under-inclusive side when imposing restrictions on industry. First, if one accepts that rule-makers tend to write over-inclusive rules for a number of reasons (as discussed above), there may be a case for countering this tendency by consciously erring towards under-inclusiveness in particular cases of uncertainty.[23] Second, one might put a value on economic liberty as a good in itself or favour under-inclusiveness where compliance costs are liable to be extremely high and the benefits of a rule are low (a regulation might be proposed in such circumstances for reasons of social justice rather than on efficiency grounds). Third, if the rule at issue exerts control at the stage of preventing a dangerous action occurring (e.g. by licensing an activity) rather than at the stage where the dangerous action has occurred or the harm has been caused,[24] there may be a case for erring on the side of under-inclusion if the costs of prevention are liable to be very high and if any problems of under-inclusiveness can be countered by controls at the act or harm stage.

To summarize, rule-making does affect both the way that enforcement is carried out and the effectiveness of enforcement activity. It can be seen also that rule-making involves complex trade-offs between, amongst other things,

[23] See Brennan, 'Civil Disaster Management', 33.
[24] On stages of intervention, see the following section of this chapter.

attempts to solve problems of inclusivity; efforts to contain creative compliers; and endeavours to produce rules that can be enforced effectively in the field.[25]

PREDICTING COMPLIANCE: THE TABLE OF ELEVEN

In considering the design of rule systems, it is worth noting a framework that has been developed by the Dutch Ministry of Justice as an aid to estimating the levels of compliance that will be associated with a given law.[26] The 'Table of Eleven' (opposite) seeks to identify the factors that impact on compliance and is highly influential in some jurisdictions as a way not merely to judge future compliance but to evaluate the enforcement efforts of a regulator and to improve the design of rules. The Table of Eleven involves reference to eleven 'causes and motives' relevant to compliance with legislation. These are divided into two groups: factors related to spontaneous compliance, and those motivations that flow from external regulatory enforcement.

The 'Table of Eleven' does not constitute a single, complete, theory of compliance—or a theory of enforcement strategy—but it does usefully draw on several bodies of theory in order to point to causes of compliance. It also offers policymakers a framework that will expose areas requiring attention together with a checklist of relevant questions. Regarding the first of the above compliance factors—'Knowledge of rules'—for instance, the checklist prompts the following queries:

a. Familiarity
- Does the target group know the rules?
- Do they only need to make limited efforts to find out about the rules?
- Is the legislation not too elaborate?

b. Clarity
- Are the rules formulated in such a way that the target group can understand them easily?
- Is the target group actually capable of understanding the rules?
- Is it sufficiently clear to the target group what the rules apply to?
- Is it clear to the target group what rule applies?

Such checklisting usefully helps to locate potential areas of weakness in a law's formulation and potential responses to flaws in strategy. The 'Table of Eleven' framework also emphasizes that it is not enough to ensure optimality of the objective risks that flow from non-compliance—what count are the target groups' subjective judgements regarding such risks.

[25] On trading-off problems of inclusiveness against other factors such as costs of enforcement, see Diver, 'Optimal Precision', 74–8, and Baldwin, *Rules and Government*, 180–5.

[26] See Dutch Ministry of Justice, *The 'Table of Eleven': A Versatile Tool* (The Hague, 2004).

THE TABLE OF ELEVEN

The spontaneous compliance dimensions

1. Knowledge of rules

a. Familiarity with rules
b. Clarity of rules

Explanation: This factor concerns the familiarity with and clarity of legislation among the target group. (Unfamiliarity with the rules may result in unintentional violations, as may lack of clarity or complexity of legislation.)

2. Costs/benefits

a. Financial/economic costs and benefits
b. Intangible costs and benefits

Explanation: This concerns all financial/economic and intangible advantages and disadvantages of compliance or non-compliance with the rule(s), expressed in time, money, and effort.

3. Extent of acceptance

a. Acceptance of the policy objective
b. Acceptance of the effects of a policy

Explanation: The extent to which the policy and legislation is considered acceptable by the target group.

4. The target group's respect for authority

a. Official authority
b. Competing authority

Explanation: The extent to which the target group respects the government's authority and the extent of respect for their *own* standards and values, which may be in conflict with the government's intentions.

5. Non-official control (social control)

a. Social control
b. Horizontal supervision

Explanation: The risk, as estimated by the target group, of positive or negative sanctions on their behaviour other than by the authorities: either through informal social controls or by 'horizontal supervision'—influences exerted by the target group or professional group on their own members.

Enforcement dimensions

6. Risk of being reported

Explanation: The risk, as estimated by the target group, of a violation being detected by others than the authorities and being reported to a government body (by tip-off or complaint, for instance).

7. Risk of inspection

a. Records inspection
b. Physical inspection

Explanation: The risk, as estimated by the target group, of an inspection into compliance by the authorities (either an inspection of records or a physical inspection).

8. Risk of detection

a. Detection in a records inspection
b. Detection in a physical inspection

Explanation: The risk, as estimated by the target group, of a violation being detected in an inspection carried out by the authorities.

9. Selectivity

Explanation: The perceived risk that a violation will be detected from within a given body of inspected businesses, persons, actions, or areas (the anticipated 'hit-rate' of inspections).

10. Risk of sanction

Explanation: The risk, as estimated by the target group, of a sanction being imposed if an inspection reveals that a rule has been broken.

11. Severity of sanction

Explanation: The severity and nature of the sanction associated with the violation and any additional disadvantages of being sanctioned (e.g. losses of respect or reputation; legal costs, etc.).

Enforcement: Strategies for Applying Tools

STYLES OF ENFORCEMENT

Regulatory officials seek to gain compliance with the law not merely by resort to formal enforcement and prosecution but by using a host of informal techniques including education, advice, persuasion, and negotiation.[27] Reiss

[27] See K. Hawkins, *Environment and Enforcement* (Oxford, 1984); B.M. Hutter, *Compliance: Regulation and Environment* (Oxford, 1997); G. Richardson, A. Ogus, and P. Burrows, *Policing Pollution* (Oxford, 1988); W.G. Carson, 'Some Sociological Aspects of Strict Liability and the Enforcement of Factory Legislation' (1970) 33 *MLR* 396; R. Cranston, *Regulating Business* (London, 1979); A. Reiss, 'Selecting Strategies of Social Control over Organisational Life', in Hawkins and Thomas, *Enforcing Regulation*; N. Gunningham, 'Enforcement and Compliance Strategies' in R. Baldwin, M. Cave, and M. Lodge (eds), *The Oxford Handbook of Regulation* (Oxford, 2010); B. Morgan and K. Yeung, *An Introduction to Law and Regulation* (Cambridge, 2007), ch. 4.

has thus drawn an important distinction between 'compliance'[28] approaches to enforcement, which emphasize the use of measures falling short of prosecution in order to seek compliance with laws, and 'deterrence'[29] approaches, which are penal and use prosecutions in order to deter future infractions.

Compliance approaches can be seen as holding conformity to the law as a more central objective than deterrence approaches, which may also involve a stronger retributive dimension.[30] Within the compliance strategy, Hutter has distinguished two sub-strategies that she terms the *persuasive* and the *insistent* approaches. Both aim to secure compliance, but the persuasive approach is more accommodating. Officials educate and coax offenders into complying with the law, they explain rationales for laws and possible means of compliance, and do so in a patient, open-ended way.[31] The insistent strategy is less flexible and there are defined limits to the tolerance of officials who will increase pressures when compliance is not forthcoming within a limited period. Cultural differences may also be seen as producing differences in enforcement styles. It has been said that the American system of enforcement tends to be more adversarial, litigious, and deterrence-based than the more compliance-orientated British approach[32] but caution should be exercised in viewing national approaches as homogeneous—variations of styles may be encountered even within single agencies (on which, more below).[33]

On the relative effectiveness and desirability of compliance and deterrence approaches to enforcement, there are a number of conflicting arguments.[34] Proponents of deterrence tend to argue that compliance approaches are indicative of capture, lack of enforcement resources, and of regulator and regulatee having sufficient identification with each other (through shared experience, contacts, staff exchanges, or familiarity) as to make routine prosecution unthinkable.[35] In favour of compliance, however, it is often

[28] See Reiss, 'Selecting Strategies'. Richardson et al., *Policing Pollution*, use the term 'accommodative' for this style.

[29] Hawkins uses the term 'sanctioning' for this approach.

[30] See Hutter, *Compliance*, 15. For a study of the conformity of enforcement approaches to constitutional values, see K. Yeung, *Securing Compliance: A Principled Approach* (Oxford, 2004).

[31] This strategy is said to approximate to Braithwaite, Walker, and Grabosky's notion of the Diagnostic Inspectorate in their article 'An Enforcement Taxonomy of Regulatory Agencies' (1987) 9 *Law and Policy* 321.

[32] See, e.g., D. Vogel, *National Styles of Regulation* (Ithaca, NY, 1988) and for a USA/Sweden comparison, see S. Kelman, *Regulating America, Regulating Sweden* (Cambridge, MA, 1981).

[33] On variations within the Health and Safety Executive's inspectorates, see Baldwin, *Rules and Government*, 143, and C.D. Drake and F.B. Wright, *Law of Health and Safety* (London, 1983), 24.

[34] For a head-to-head confrontation, see the exchange between F. Pearce and S. Tombs, 'Ideology, Hegemony and Empiricism' (1990) 30 *British Journal of Criminology* 424, and also K. Hawkins, 'Compliance Strategy, Prosecution Policy and Aunt Sally' (1990) 30 *British Journal of Criminology* 144, and 'Enforcing Regulation: More of the Same from Pearce and Tombs' (1991) 31 *British Journal of Criminology* 427. See also Ayres and Braithwaite, *Responsive Regulation*, 20–35.

[35] See R. Brown, 'Theory and Practice of Regulatory Enforcement: Occupational Health and Safety Regulation in British Columbia' (1994) 16(1) *Law and Policy* 63.

argued that this approach offers an efficient, cost-conscious use of resources and can be justified as economically rational, rather than indicative of capture.[36] Prosecutions are so costly in time and money that selective use of less formal mechanisms may produce higher levels of compliance for a given level of state expenditure than is possible with routine prosecution.[37] Compliance approaches, it is further said, are better than deterrence strategy in fostering valuable information flows between regulators and regulated; in educating regulatees so that they think constructively about modes of compliance; and in encouraging firms to perform to higher standards than the law calls for.[38]

The case for deterrence approaches is, in contrast, based on claims that they can prove highly effective in changing corporate cultures and stimulating management systems that reduce the risks of infringement.[39] Deterrence approaches treat infractions seriously by stamping errant conduct as unacceptable—they accordingly reinforce, and give effect to, social sentiments of disapproval, and this enhances social pressures to comply. Such approaches, it is urged moreover, can induce political shifts so that firmer approaches to regulation can be taken.[40] Tough approaches to enforcement, say deterrence advocates, make it rational for firms to give a high priority to compliance.

There are, however, a number of reasons to question the effectiveness of deterrence approaches in controlling corporations. Deterrence strategy demands that sanctions are sufficiently severe to incentivize compliance beyond the immediate case. In many instances, however, the courts do not impose fines of sufficient gravity to produce this result.[41] This may happen for a number of reasons—a large fine may be seen as threatening an errant company's solvency and employees' livelihoods, for instance, or it may be viewed as depriving the rule breaker of the liquidity that is needed to remedy the mischief at issue. Whatever the cause, the frequent effect is under-inducement to comply.[42]

The assumptions of rationality that underpin deterrence theory are also questioned by numbers of critics. As with individuals, instances of non-compliance within companies may be the result of 'irrationalities' that flow

[36] See P. Fenn and C. Veljanovski, 'A Positive Economic Theory of Regulatory Enforcement' (1988) 98 *Economic Journal* 1055; C.G. Veljanovski, 'Regulatory Enforcement: An Economic Study of the British Factory Inspectorate' (1983) 5 *Law and Policy Quarterly* 75; Hawkins, *Environment and Enforcement*, 198–202.

[37] See Baldwin, *Rules and Government*, ch. 6.

[38] Hawkins, *Environment and Enforcement.*

[39] See Bardach and Kagan, *Going by the Book*, 93–5.

[40] See Pearce and Tombs, 'Ideology, Hegemony and Empiricism', 434.

[41] See R. Macrory, *Regulatory Justice: Sanctioning in a Post-Hampton World* (Cabinet Office, May 2006).

[42] For doubts whether increased fines necessarily increase compliance incentives, see S. Kadambe and K. Segerson, 'On the Role of Fines as an Environmental Enforcement Tool' (1998) 41 *Journal of Environmental Planning and Management* 217–26.

from such sources as poor training and ill-organization. Responding to such problems with the tool of rational deterrence may, accordingly, involve a mismatch of response and mischief.[43] For such and other reasons, it is not always certain whether punitive sanctions will work to stimulate internal controls.[44] It is worth emphasizing that corporations will often be confused and irrational about their risks of being punished for non-compliance (their 'punitive risks'); their staff may conflate individual and corporate liabilities; they may be poorly organized to deal with, anticipate, or react to punitive risks and the effects of sanctions; their Boards may under-perform in supervising or providing leadership on punitive risk-management, and they may be poorly placed to assess how they and their staff are performing as risk managers.[45] Many small and medium companies, indeed, will not know what it would cost to comply with the rules, and they may tend to assume that they are compliant until they are told otherwise.[46] They may, moreover, see compliance not as observing the rules but as behaving in accord with the last negotiation that they had with an inspector.[47] In the face of such propensities, it is difficult to see deterrent forces as bearing materially on regulated concerns.

Even if it is assumed that some companies do act along rational lines, it cannot be taken for granted that, when they do react to possible punitive liabilities, this will lead them to respond with compliance. Even on rational deterrence assumptions, they may see compliance as just one way to reduce the probability of suffering a sanction for breaching the rules. Other ways to reduce that probability might include creative compliance, bringing pressure to bear on the regulator (to discourage prosecution); shifting risk or blame on to the shoulders of individuals or outsourced business partners; or evasion, non-cooperation with regulators, and concealment.[48]

[43] Ibid. and see S. Stafford, 'The Effect of Punishment on Firm Compliance with Hazardous Waste Regulations' (2002) 44 *Journal of Environmental Economics and Management* 290–308.

[44] See B. Fisse and J. Braithwaite, *Corporations, Crime and Accountability* (Cambridge, 1993); C. Parker, *The Open Corporation* (Cambridge, 2002); N. Gunningham and R. Johnstone, *Regulating Workplace Safety* (Oxford, 1999), 217–25; C.D. Stone, *Where the Law Ends: The Social Control of Corporate Behaviour* (New York, 1975); J. Braithwaite and T. Makkai, 'Testing an Expected Utility Model of Corporate Deviance' (1991) 25 *Law and Society Review* 7; R. Brown, 'Administrative and Criminal Penalties in the Enforcement of the Occupational Health and Safety Legislation' (1992) 30(3) *Osgoode Hall Law Journal* 691. On deterrence in the environmental sector, see D. Chappell and J. Norberry, 'Deterring Polluters: The Search for Effective Strategies' (1990) 13 *University of New South Wales Law Journal* 97.

[45] R. Baldwin, 'The New Punitive Regulation' (2004) 67 *Modern Law Review* 351–83.

[46] J. Braithwaite and T. Makkai, 'Testing an Expected Utility Model of Corporate Deterrence' (1991) 25 *Law and Society Review* 7–39; H. Genn, 'Business Responses to the Regulation of Health and Safety in England' (1993) 15 *Law and Policy* 219–33.

[47] See R. Fairman and C. Yapp, 'Enforced Self-Regulation, Prescription and Conceptions of Compliance within Small Businesses' (2005) 27 *Law and Policy* 491–519.

[48] On the management of liability risks, see W. S. Laufer, 'Corporate Liability Risk Shifting and the Paradox of Compliance' (1999) 52 *Vanderbilt Law Review* 1343, 1410–11; J. Black, 'Using Rules

The prospect of possible sanctions may also induce companies to respond by reducing the impact of any sanction. They may thus take out insurance, share risks, or act to increase corporate resilience. On rational deterrence assumptions, companies will balance the costs and benefits of compliance with the expected costs and benefits of non-compliance, and will choose the combinations of risk-reducing responses that maximize benefits over costs.[49]

The predominant position is, however, that many companies operate largely unaware of their exposure to punitive regulatory risks, and the overall evidence is not highly consistent with effective and rational regimes in which anticipated penalties stimulate compliance. It is, however, in tune with the findings of those researchers who argue that: 'The bounded rationality of organizations and top management means that many do not make rational cost-benefit calculations about compliance until something happens to bring the risks of non-compliance to their attention. Economic costs which do not draw attention to themselves by generating some kind of crisis are often overlooked by busy management.'[50]

Other posited difficulties with deterrence approaches are that they can often prove inflexible; they may fail to identify the best ways to improve performance; and they may cause resentment, hostility, and lack of cooperation in those regulated. Where such resistance is forthcoming from regulated parties, this, in turn, is likely to reduce the effectiveness of enforcement, as well as increase overall costs.[51] A further criticism of deterrence-based enforcement is that it may produce undesired side-effects. Thus, as noted, it may drive certain firms out of business and cause unemployment, with the result that the public is alienated and, in turn, there is a reduction in regulatory effectiveness. Compliance strategies, on the other hand, can provide responses to risks (as opposed to harms actually inflicted) and can do so more effectively than deterrence approaches. They can thus prevent more harms from occurring.[52] They are consistent with making exceptions where there is a

Effectively' in C. McCrudden (ed.), *Regulation and Deregulation* (Oxford: 1999) 118–19; R. Kraakman, 'Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy' (1986) 2 *Journal of Law, Economics and Organization* 53; and S. S. Simpson, *Corporate Crime, Law and Social Control* (Cambridge, 2002), 51–2. On creative compliance, see D. McBarnet and C. Whelan, 'The Elusive Spirit of the Law: Formalism and the Struggle for Legal Control' (1991) 54 *Modern Law Review* 848. For a discussion of risk-shifting by moving business activity to less rigorous legal environments, see C.E. Moore, 'Taming the Giant Corporation' (1987) 33 *Crime and Delinquency* 388, 390–4.

[49] A further barrier to compliance is noted by Diane Vaughan. In processes of 'deviance normalization' corporate decisions, practices, and cultures may stretch the boundaries of acceptable behaviour beyond the legal so that non-compliance is routinized—see D. Vaughan, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA* (Chicago, 1996).

[50] Parker, *The Open Corporation*, 69–70.

[51] See R. Kagan and J. Scholz, 'The Criminology of the Corporation and Regulatory Enforcement Strategies', in Hawkins and Thomas, *Enforcing Regulation*.

[52] See, e.g., R. Brown and M. Rankin, 'Persuasion, Penalties and Prosecution', in M.L. Friedland (ed.), *Securing Compliance* (Toronto, 1990).

case for them—rather than imposing uniform requirements where this makes no sense.[53]

Instead of drawing a stark contrast between compliance and deterrence approaches to enforcement it is possible, however, to see enforcement as involving a progression through different compliance-seeking strategies and sanctions. Ian Ayres and John Braithwaite comment: 'To reject punitive regulation is naive; to be totally committed to it is to lead a charge of the light brigade. The trick of successful regulation is to establish a synergy between punishment and persuasion.'[54]

These two authors popularized the concept of enforcement pyramids. One of these pyramids involves a hierarchy of sanctions, the second, a hierarchy of regulatory strategies. In this model of 'responsive regulation', those regulated are subjected to increasingly interventionist regulatory responses as they continue to infringe, and to less interventionist actions as they come to comply. The 'responsive regulation' approach is considered in the following chapter, as are a number of other enforcement theories that constitute alternatives to the traditional choice between deterrence and compliance.

WHEN TO INTERVENE

Regulators can intervene in economic or social activity not merely by different methods but at different stages in the processes that lead to harms. Thus, action can be taken to *prevent* a dangerous act or situation arising (e.g. hotels can be inspected and licensed before opening to guests in order to ensure that fire hazards do not arise); action can be taken in response to the *act* of creating a dangerous situation (e.g. operating a hotel without fire doors or sprinklers); or action can be prompted by the realization of a *harm* (e.g. injuring a hotel guest in a hotel fire). Similarly, the kinds of standards incorporated in regulatory rules may be directed towards different stages of the processes leading to harms. Thus, applying *specification* (or *design*) standards (e.g. demanding a certain type of machine be used for a procedure) has the effect of controlling the circumstances leading to dangerous acts (which will not occur if a safe design of machine is used). Applying *performance* (or *output*) standards looks to the seriousness of the dangerous acts that are involved in a procedure, and *target* standards focus on the harms that result (e.g. the amount of pollution damage to the river).[55]

[53] See Bardach and Kagan, *Going by the Book*, 58.

[54] Ayres and Braithwaite, *Responsive Regulation*, 25; see also J. Braithwaite, *To Punish or Persuade* (Albany, NY, 1985).

[55] For further discussion of standard-setting, see Chapter 14.

Shavell has distinguished three stages of intervention, in which actions are respectively preventative, act-based, and harm-based, and has considered the circumstances that favour particular stages of regulatory intervention.[56]

PREVENTATIVE ACTIONS

One circumstance in which preventive action is called for is where the costs of rectifying a dangerous state of affairs may be high. It may thus be better to prevent a dangerous design of steel foundry from being built than to try and change matters post-construction. Another reason for acting preventively is that a potentially catastrophic danger may otherwise arise. For both of these reasons, nuclear reactor designs are approved and licensed before the reactor is built. To allow a reactor to be constructed and then to demand changes to improve safety would be hugely expensive and to intervene only when a reactor was being run dangerously would be to run an unacceptably high risk of a catastrophe occurring. Conversely, where the potential harm is relatively small, there may be a stronger case for intervening at the harm stage, since failures to deter through punishing harm-causers will not prove catastrophic.

A further reason for preventative action is that in some circumstances, preventative steps can be successfully employed without the use of a heavy regulatory hand. Thus, a television advertisement on public safety (for example on dipping car headlights properly when night driving) may prevent a large number of dangerous acts and serious harms from arising, but it neither interferes greatly with drivers nor involves strong sanctions or expensive enforcement activity. By comparison, intervening at the act stage—by stopping drivers and bringing legal proceedings for failing to use lights properly—involves a good deal of driver inconvenience and state enforcement expense. It may also fail to prevent a large number of harms from ensuing. Waiting until the harms occur (the injuries and deaths) will demand expensive prosecutions and severe sanctions and such actions may (for reasons discussed below) also fail efficiently to prevent future harms from occurring.

The potential of preventative strategies is often restricted by informational difficulties. It may not be easy to predict when certain dangerous situations or harms may occur. To prevent dangerous driving manœuvres would thus be difficult and costly. Even if all drivers had an inspector sitting in their vehicle it would be difficult for the inspector to anticipate when a dangerous turn or stop was about to happen.[57] It may, accordingly, be more efficient to exert

[56] See S. Shavell, 'The Optimal Structure of Law Enforcement' (1993) *Journal of Law and Economics* 255.

[57] See Shavell, 'The Optimal Structure of Law Enforcement', 272. The licensing and testing of drivers does, however, constitute prevention at the most general level.

control at the act or harm stage by imposing sanctions for the act of dangerous driving or for causing an accident. It may, similarly, be difficult for regulators to anticipate all potential sources of harm (e.g. of damaging levels of noise in industrial processes). Preventing dangerous activity in such circumstances may, accordingly, make very severe informational demands and prove very expensive. In terms of administrative costs it will often be cheapest of all for the state to wait until the stage when harms have been caused, since only a small proportion of dangerous acts will result in actual harms and so a smaller enforcement case load will be involved.

ACT-BASED INTERVENTIONS

Intervening at the act stage may be more useful than at the preventative stage when prevention would be expensive to accomplish. Thus it might be costly and intrusive to demand prior approval for all operations on a construction site. Resources are more effectively deployed in inspecting such sites and sanctioning dangerous actions—for example, using insecure scaffolding.

Act-stage intervention may be preferable to harm-stage control where it is difficult to hold firms or individuals to account for causing harms. This may well be the case where a harm may arise from a number of concurrent sources. Thus, if it is known that ingesting certain particles may cause cancer but it is also known that the disease can be caused by a number of other common hazards (e.g. smoking), it may be preferable to regulate, say, the act of causing workers to ingest the particles than to look at the harm and attempt to establish that a particular cancer has been caused by the particles rather than some other agent.

Where the dangerous act can be identified far more easily than the resultant harm, there is, again, a case for act-based intervention. The harms caused by some acts may be cumulatively very serious but highly diffused (e.g. in the pollution field), and it will be more effective to enforce at the act level then to deal with huge numbers of individually small claims.

Intervening at the act stage may also be preferable to harm-based sanctioning where the latter will under-deter the causing of further harms. Act-based sanctions can be smaller and more frequently applied than harm-based ones because no injury has yet occurred and the instances of infraction are greater in number (since not all dangerous acts lead to harms). In order to obtain the same levels of deterrence from harm-based sanctions, the fines involved may have to be extremely high since a hazard creator may estimate the chance of actually producing a harm as extremely small and of being fined as even smaller. If the firms causing harms are unlikely to be able to pay such fines, they will be under-deterred. In some sectors there will be a correlation between poor resourcing and the causing of harms and, accordingly, there

will tend to be particular problems of under-deterrence through harm-based intervention. Where potential harms may vastly exceed the resources of the causers of harm (e.g. where ill-resourced firms may injure or kill large numbers of persons) there is liable to be under-deterrence and unacceptably low incentives to take appropriate care.

HARM-BASED INTERVENTIONS

Enforcement costs may sometimes militate in favour of intervening at the stage of harm, rather than of dangerous act. It may be cheaper for the state to punish the causers of the relatively small number of instances of harm than to pursue the much larger number of persons who perform dangerous acts that are liable to cause harms in some instances. It follows from what has been said above that a policy of severely sanctioning harms will only deter adequately if detection is sufficiently effective and if potential offenders are likely to be able to pay the large fines involved (or serve the prison terms).

It cannot, however, be assumed that in all cases it is easier to apply sanctions to harms than to acts. If a construction firm erects cheap but unsafe scaffolding and makes it difficult for a regulator to detect harms—for example, by rewarding injury-free teams of workers with bonuses and thereby generating peer group pressures on workers not to report injuries—it may be easier for a regulator to penalize the (highly visible) act of using dangerous scaffolding than to punish the firm for occasioning harm to the worker.

Enforcement is a matter of deploying a strategy or mixture of targeted strategies for securing desired results on the ground.[58] The NAO found that Defra fisheries regulators prioritized inspections according to broad-based risk analyses which tended to target particular fisheries and types of activity, rather than individual vessels.[59] Thus, surveillance operations and inspections tended to focus on areas of high risk where quotas were most restrictive, stocks were of high value, fishing activity was intense, fish were known to be collecting, or fisheries were seasonal. Inspections also tended to be concentrated on points where the regulatory returns to interventions would have been the greatest—for example on those ports where landings were, given the circumstances, most likely. Major difficulties encountered in using such risk-based approaches lay in coming to terms with new risk creators and new risks

[58] See generally, Hutter, *Compliance*, ch. 1; id., *The Reasonable Arm of the Law?* (Oxford, 1988); Hawkins, *Environment and Enforcement*. On private enforcement, see J. Braithwaite, 'Enforced Self-Regulation' (1982) 80 *Michigan Law Review* 1461; Ayres and Braithwaite, *Responsive Regulation*, ch. 4; C.D. Shearing and P.D. Stenning (eds), *Private Policing* (Beverly Hills, 1986); W. Landes and R. Posner, 'The Private Enforcement of Law' (1975) 4 *Journal of Legal Studies* 1.

[59] See National Audit Office (NAO), *Fisheries Enforcement in England*, HC 563 Session 2002–3 (April 2003) (hereafter 'NAO, *Fisheries Enforcement*'), 20. On risk-based regulation, see chapter 13 below.

to fish stocks. The extent of 'off screen' activities also tended to undermine the reliability of the data underpinning the risk analyses.

HOW MUCH TO ENFORCE

It is not sensible for regulators to aim for perfect compliance or the complete elimination of a risk. This is because enforcement costs tend to escalate as targeted levels of compliance are raised, and a point will arrive where the costs of further enforcement are not justified by the gains. Breyer refers to this as the problem of the last 10 per cent[60] and quotes Sheldon Meyers: 'it frequently is relatively cheap to reduce risks from 0 to 90 per cent, more expensive to go from 90 per cent to 99 per cent and more expensive to go from 99 per cent to 99.9 per cent'.[61]

In economic terms, the socially optimal level of enforcement occurs at the point where the extra costs of enforcement exceed the resulting additional benefits to society.[62] Included within the costs of enforcement are the following:

- the costs of agency monitoring;
- the expenses of processing and prosecuting cases;
- the defence costs (of innocent and guilty parties);
- the costs of misapplications of law, convicting the innocent, and deterring desirable behaviour.

The gains from enforcement lie principally in reductions of harmful behaviour—be this from preventing the particular offender from causing harm or from deterring others. A further gain, however, flows from reductions in private enforcement costs. Thus, when public enforcement agencies forestall a harm, this saves private individuals or firms from having to spend money on protecting their entitlements.

In calculating the deterrent effects of enforcement activity, the economic approach assumes *inter alia* that potential offenders are actors who seek to maximize their own welfare in an informed, rational manner. For each potential offender deterrence flows from the expected punishment, which is

[60] S. Breyer, *Breaking the Vicious Circle* (Cambridge, MA, 1993), 10–13.

[61] S. Meyers, 'Applications of *De Minimis*', in C. Whipple (ed.), *De Minimis Risk* (New York, 1987), 102.

[62] See generally, G. Becker, 'Crime and Punishment: An Economic Approach' (1968) 76 *Journal of Political Economics* 161; I. Ehrlich, 'The Economic Approach to Crime—a Preliminary Assessment', in S. Messinger and E. Bittner, *Criminology Yearbook* (London, 1979); Ogus, *Regulation*, 90–4; Shavell, 'The Optimal Structure of Law Enforcement'; G.J. Stigler, 'The Optimum Enforcement of Laws' (1970) 78 *Journal of Political Economics* 526; T. Gibbons, 'The Utility of Economic Analysis of Crime' (1982) 2 *International Review of Law and Economics* 173.

the probability of punishment times the magnitude of the punishment (e.g. the quantum of the fine).

From the regulator's point of view, a key calculation is how much the group of potential offenders will be deterred by the regulator's current or prospective approach to enforcement.[63] Factors to be taken on board include not merely the level of fines or other sanctions liable to be applied and the probability of inflicting these on offenders, but also the private benefit likely to be derived from offending and the social cost of the offence. The overall wealth of the offender has also to be considered. If an offending firm cannot pay a large fine (because, say, this would drive the enterprise into insolvency and cause unemployment), a combination of small fines and high probability of application would be more appropriate than using large penalties infrequently. Similarly, if severe sanctions are unlikely to be applied for reasons of social justice (the courts may consider the offence minor) a high probability of application will have to be used, especially if the gains from offending are high.

Where, on the other hand, enforcement resources are limited and the probability of bringing sanctions to bear is, as a result, low, it may be rational for the regulator to press the appropriate authorities for penalties great enough to compensate for this improbability.[64] In response to arguments that fairness imposes limitations on the quantum of a punishment for a given offence, the regulator may reply that what matters in real life is the *expected* punishment—that if governments want low-resource regulation, they have to be prepared to impose high penalties.

To balance such talk of economic rationality, it should first be noted that policy and equitable considerations may often govern enforcement decisions. Thus, as a matter of policy, society may want to deter certain activities very strongly and not rely on an efficiency-based balancing of expected gains and penalties. Second, the assumptions of economically rational man may be questioned. In the real world, most harms are not the result of rational calculations concerning costs and benefits—they are the products *inter alia* of human failings, poor information and training, tiredness, short cuts, and accidents.[65] In so far as the model of rational man fails accurately to describe those persons or firms that are regulated, the regulator may feel (and be)

[63] On the imperfections of the expected cost approach to deterrence, see T. Makkai and J. Braithwaite, 'The Limits of the Economic Analysis of Regulation: An Empirical Case and a Case for Empiricism' (1993) 15(4) *Law and Policy* 271.

[64] In September 1997, the Chief Executive of the Environment Agency made a strong public attack on the current level of fines for environmental offences. During the same month, Michael Meacher, the Environment Minister, announced that the Government was drawing up plans for large increases in fines for persistent corporate polluters; *Financial Times*, 17 Sept. 1997.

[65] Makkai and Braithwaite, 'The Limits of the Economic Analysis of Regulation'.

justified in placing less emphasis on deterrence and more on active inspection and intervention in the regulated activity.

CONTROLLING CORPORATIONS

Much of regulation concerns the control of corporations, and it is worth noting here the particular difficulties that are likely to be encountered in seeking to control errant companies. Key issues concern the sanctions that can be used to influence such firms; the extent of corporate criminal fault; and the difficulties of proving liability.[66]

SANCTIONS

Regulators can resort to administrative or criminal sanctions in dealing with corporations and, as noted, a wide variety of instruments can be applied. Administrative measures operate without recourse to the courts and can be provided for either in statutes or in contracts (e.g. within the terms of franchises). Examples of statutory administrative sanctions include improvement and prohibition notices, which respectively require remedial actions to be taken within a fixed period or which order the discontinuance of a hazardous activity (e.g. the stopping of a dangerous production line).[67] Contractually based measures may include licence revocations, curtailments, or suspensions.

Criminal sanctions normally involve fines since imprisoning firms is not feasible—though directors may be found criminally liable as individuals, for example, where their personal gross negligence has resulted in a death.[68] To impose fines on firms that, say, pollute waterways or impose health risks on employees, can, however, give rise to difficulties. The firm may engage in activities liable to cause harms that have a value that exceeds any fine they are

[66] On criminalizing corporations, see generally: C. Wells, *Corporations and Criminal Responsibility* (Oxford, 1993); L.H. Leigh, *The Criminal Liability of Corporations in English Law* (London, 1969); id., 'The Criminal Liability of Corporations and Other Groups: A Comparative View' (1982) 80 *Michigan Law Review*; J.C. Coffee, 'No Soul to Damn, No Body to Kick' (1981) 79 *Michigan Law Review* 386; C. Clarkson, 'Kicking Corporate Bodies and Damning their Souls' (1996) 59 *MLR* 557; B. Fisse and J. Braithwaite, 'The Allocation of Responsibility for Corporate Crime' (1988) 11 *Sydney Law Review* 468; C. Stone, 'The Place of Enterprise Liability in the Control of Corporate Conduct' (1980) 90 *Yale LJ* 1; R.A. Kagan and J.T. Scholz, 'The Criminology of the Corporation and Regulatory Enforcement Strategies', in Hawkins and Thomas, *Enforcing Regulation*; T. Kaye, 'Corporate Manslaughter: Who Pays the Ferryman?', in D. Feldman and F. Meisel (eds), *Corporate Commercial Law: Modern Developments* (London, 1996).

[67] See Health and Safety at Work etc. Act 1984, ss. 21, 22.

[68] Following the Lyme Bay canoeing tragedy in which four teenagers were drowned, both the managing director of the company that owned the responsible outdoor activity centre and the company itself were found guilty of manslaughter; see Kite and Others, in *The Independent*, 9 Dec. 1994.

able to pay. Any potential fine will accordingly under-deter.[69] Firms, moreover, may treat fines as a normal business expense and may be able to pass the cost of fines through to consumers or even employees.[70] Large fines may prejudice the firm's survival, and insolvency may punish innocent parties such as employees or customers.

Fines remove ready cash from the company which might have been spent on measures to limit the harms at issue (e.g. on new filtration systems to reduce pollution). Fines that do come to bear on the corporation may, however, not deter or influence the actual decision-maker within the management structure,[71] and fines do not ensure that the problem at issue will be remedied or that the causes of failure within the corporation will be identified.

Alternative ways of sanctioning corporations have been suggested in an attempt to improve on the deficiencies of fines. A first of these is the equity fine.[72] Under an equity fine system, the convicted corporation is required to issue a given number of shares to the state's victim compensation fund. The shares would have a value equivalent to the cash fine necessary to deter the illegal activity. This strategy has the supposed advantages that it reduces the negative effect of corporate penalties on workers and consumers since the costs of deterrence are concentrated on the shareholders (whose shares lose value as a result of the mandated issue). These shareholders will accordingly have an incentive to discipline managers. The threat of insolvency and harm to employees and the community is removed. High penalties can be imposed because the market valuation of the typical corporation vastly exceeds the cash resources available to it (cash that would be the target of any fines imposed). Cash is not removed from the corporation, and so spending on harm avoidance is not prejudiced. Managers' interests are aligned with those of the corporation in so far as stock options will lose value on a mandated issue. Mandated issues will produce managerial fears of takeovers, and this will provide an incentive to good behaviour on the part of managers and, finally, shareholders will demand internal controls to reduce dangers of stock dilution through mandated issues. These controls will help avoid regulatory infringements.

Equity fines are thus superior to fines in a number of respects, but they may not prove popular with governments that are opposed to state equity holdings and they share, with fines, the weakness that their deterrent value depends on

[69] See Coffee, 'No Soul to Damn, No Body to Kick', 389–93.

[70] In response, it can be argued that consumers should pay a price for goods that reflects the costs of production (which should include any social costs imposed e.g. by pollution); see B. Fisse, 'Sentencing Options against Corporations' (1990) *Criminal Law Forum* 211; C. Stone, 'Controlling Corporate Misconduct' (1977) *Public Interest* 55; Coffee, 'No Soul to Damn, No Body to Kick'.

[71] See Fisse and Braithwaite, 'Allocation of Responsibility'.

[72] See Coffee, 'No Soul to Damn, No Body to Kick', 413–15.

the probability of apprehension and punishment. They furthermore make little contribution to the reform of the corporation's internal procedures and do not ensure that guilty managerial parties will be disciplined.[73]

Another, often-proposed, alternative to the fine is the punitive injunction.[74] Courts could use punitive injunctions to require corporations to remedy their internal controls and to introduce (perhaps at punitive expense) preventative equipment or procedures. Further devices are corporate probation orders[75] and enforced accountability regimes.[76] Judges who are empowered to institute such orders and regimes are able to monitor the activities of a convicted organization and to insist on reporting, record-keeping, and auditing mechanisms that are designed to remedy identified failings and to hold individuals to account. Corporations can be ordered to undertake enquiries, apply discipline, and report on steps taken, and senior managers can be threatened with personal criminal liability if they fail to take such steps to the satisfaction of the court. As with punitive injunctions, particular errant managers or sections of management can be identified and their deficiencies addressed.

Courts can also be given powers to make community service and compensation orders. These can compel corporations to provide certain services for the community or to compensate individuals or groups in an attempt both to make good harms done and to signify the need for corporate rehabilitation. Adverse publicity orders can, in addition, be deployed to instruct corporations to place notices in the media informing the public of their failings and of remedial measures taken.

All such devices have their strengths, weaknesses, and areas of most useful application. It is perhaps appropriate, therefore, for regulators and courts to approach corporate failure with the full array of such sanctions within their contemplation and to apply them bearing in mind not merely the need to punish and rehabilitate corporations but also the interests of the public in compensation, where appropriate, and in more effective compliance.

THE EXTENT OF CORPORATE FAULT AND PROVING LIABILITY

For many years, the regulators of corporations faced a significant legal challenge in Britain. It was difficult to attribute criminal responsibility to corporations because the criminal law had developed with an eye to individual fault, and for liability to apply to a corporation, it was necessary to identify an

[73] See Wells, *Corporations*, 35.

[74] See Fisse and Braithwaite, 'Allocation of Responsibility', 500.

[75] See Coffee, 'No Soul to Damn, No Body to Kick', 448–57; Stone, 'Controlling Corporate Misconduct'; Wells, *Corporations*, 36–7.

[76] See Fisse and Braithwaite, 'Allocation of Responsibility'.

individual managerial representative of the company who had been blameworthy: who had carried out the prohibited act (*actus reus*) with the guilty mind (*mens rea*) that the relevant offence required.[77] The deficiencies of this 'identification' approach were exposed following the P&O case of 1990,[78] which arose from the deaths of 187 people in the capsize of the *Herald of Free Enterprise* after it set sail with open doors at Zeebrugge in 1987. Acquittals were directed in the case of the P&O company and its five most senior employees, since it could not be proved that the risks of open-door sailings were obvious to any of the senior managers. As a result, no *mens rea* could be attributed to the company.[79]

In the mid-nineties, however, a less restrictive view of corporate criminal liability was taken in the courts. The Privy Council, in the *Meridian* case,[80] rejected exclusive reliance on the identification test and indicated that acts and knowledge can be attributed to a company by courts considering whose acts, knowledge, or state of mind was *for the purpose* of a particular law to count as belonging to the company. Thus, instead of applying a simple identification test, the judges would, in such an approach, consider the language of a rule, its content and policy, and construe corporate liability accordingly. The functions actually performed by individuals in the company become relevant, rather than their status in the company hierarchy—a mode of reasoning liable to lead to corporate responsibility for the acts of those at lower levels than would be the case under exclusive reliance on the identification principle.

The restrictiveness of the *Tesco* identification doctrine was also circumvented by a different route—that of vicarious liability.[81] Thus in *National Rivers Authority* v. *Alfred McAlpine Homes East* [1994] CLR 760, two employees were responsible for allowing wet cement to pollute a controlled water, but, at trial, their employing company was acquitted of the statutory pollution offence under the identification doctrine of *Tesco* v. *Nattrass*. On appeal, however, the Divisional Court applied the doctrine of vicarious liability. The court looked at the purpose of the pollution legislation, bore in mind that the

[77] See *Tesco Supermarkets Ltd* v. *Nattrass* [1972] AC 153; *H.L. Bolton (Engineering) Co. Ltd* v. *T.J. Graham and Sons Ltd* [1957] 1 QB 159. See generally C. Wells and O. Quick, *Lacey, Wells and Quick: Reconstructing Criminal Law* (4th edn. Cambridge, 2010).

[78] *R* v. *Alcindor and others* (Central Criminal Court, 19 Oct. 1990); *R* v. *P&O European Ferries (Dover) Ltd* (1980) 83 GAPP. R. 72; *P&O European Ferries Ltd* (1991) 93 Crim. App. R. 72; *R* v. *Coroner for East Kent ex. p. Spooner* (1989) 88 Crim. App. R. 10.

[79] Clarkson, 'Kicking Corporate Bodies', 561. Inquiries found companies to be seriously at fault, but no successful prosecutions for manslaughter were brought, following the 1987 King's Cross fire (in which 31 died); the 1988 Clapham rail crash (in which 35 died); and the 1988 Piper Alpha oil platform disaster (in which 167 died).

[80] *Meridian Global Funds Management Asia Ltd* v. *The Securities Commission* [1995] 3 WLR 413; also R. Grantham, 'Corporate Knowledge: Identification or Attribution?' (1996) 59 *MLR* 732.

[81] See Clarkson, 'Kicking Corporate Bodies', 563–6.

offence was one of strict liability (that is, it did not require proof of a guilty mind, only that the accused caused the prohibited action), and held that the nature of the offence demanded that vicarious liability be imposed on the company for the acts of employees (whether they represented the directing mind and will of the company or not). This approach has been followed in the Court of Appeal[82] and in cases where proof of negligence has been required,[83] though the House of Lords has cautioned that (at least under the Merchant Shipping Act 1988) a company could not be held liable for each and every wrongful act committed by any employee.[84]

The problems left unsolved by the *Meridian* approach were, first, that it was not wholly clear when the action of a person who did not represent the directing mind and will of the company would be attributed to the company.[85] Second, it was still necessary to find some person within the company who had perpetrated the criminal acts yet, in real life, regulatory failings may (as in the P&O case) stem from general managerial slackness and failures to allocate responsibilities rather than from the identifiable actions of particular individuals.

With respect to manslaughter by corporations, a long period of campaigning, reviews, and debates on law reform finally led to the eclipsing of the *Tesco* doctrine by the passing of the Corporate Manslaughter and Corporate Homicide Act 2007.[86] That legislation stipulated that a corporation will be liable for the offence of corporate manslaughter if the way in which its activities are managed or organized: (a) causes a person's death; and (b) amounts to a gross breach of a relevant duty of care owed by the organization to the deceased.[87] Liability only occurs, however, if the way in which its activities are managed or organized by senior management is a substantial element in the breach of the duty of care[88] and a breach is only 'gross' if the relevant conduct

[82] See *R* v. *British Steel Plc* [1995] 586.

[83] *Tesco Stores Ltd* v. *Brent LBC* [1993] 2 All ER 718; *Re: Supply of Ready Mixed Concrete* (No. 2.) [1995] 1 All 135.

[84] *Seabound Offshore Ltd* v. *Secretary of State for Transport* [1994] 2 All ER 99.

[85] See Clarkson, 'Kicking Corporate Bodies', 566.

[86] See, for example: Law. Com. No. 237, *Legislating the Criminal Code: Involuntary Manslaughter* (1996), paras 8.1–8.77 and Draft Involuntary Manslaughter Bill s. 4 (1). In October 1997, Home Secretary Jack Straw argued in favour of a 'corporate killing' offence at the Labour Party Conference. See also TUC, *Paying the Price for Deaths at Work* (London, 1994); Wells, *Corporations*, 144–5.

[87] Corporate Manslaughter and Corporate Homicide Act 2007 s.1. Under section 2 of the Act, a duty of care is owed for the purposes of the 2007 Act if, *inter alia*, a duty is owed to employees under the law of negligence or a duty is owed as occupier of premises, supplier of goods and services, or the carrying on of any construction, maintenance, or other commercial activity. On the Act, see J. Gobert, 'The Corporate Manslaughter and Corporate Homicide Act 2007' (2008) 71 *MLR* 413 and D. Ormerod and R. Taylor, 'The Corporate Manslaughter and Corporate Homicide Act 2007' (2008) *Criminal Law Review* 589.

[88] Corporate Manslaughter and Corporate Homicide Act 2007 s.1(3).

falls 'far below what can reasonably be expected of the organization in the circumstances'.[89]

Regulatory enforcement officials, as can be seen, face not inconsiderable legal difficulties in attempting to hold corporations to account by means of the criminal law. It might, moreover, be asked: 'Why, in any event, punish corporations criminally?'[90]

A first reason is that community disapproval calls, in some instances, for the stamp of criminalization to be imposed. A second is that use of the criminal law provides a set of useful incentives that are of value even when there is no great need to stigmatize conduct as particularly heinous. A third is that corporations, just like individuals, can make decisions and have the capacity to change their policies and procedures and accordingly do meet the conditions of blameworthiness and responsibility.[91] They can, moreover, be deterred by threats of punishment. Finally, the corporation may be better placed than the state to put right its internal failings, and so it may be sensible to use the criminal law to give the firm an incentive to do this. The corporation, moreover, is more likely (for informational and evidential reasons) to apply sanctions to an errant manager than is the state, and the ensuing higher 'expected punishment cost' that flows from internal controls means that higher levels of deterrence may be attained by punishing the corporation, and leaving it to take further action, than are secured by focusing the criminal law directly on the errant employee.

Assessment

A fourth task within regulatory enforcement is the development of performance sensitivity through processes that not only evaluate how well the current system is being enforced, but also calculate how much undesirable activity is escaping the impact of the current regime of controls. This task involves assessing the strength of the case for developing new tools, or adopting new enforcement strategies, or moving towards a new design of regulatory regime. Performance assessment is thus centrally important for the progressive development of regulatory policies and is integral to good regulatory management—especially across complex networks of state and other controls. It is also essential to accountability and transparency insofar as assessments provide measures of progress in meeting objectives and their

[89] Corporate Manslaughter and Corporate Homicide Act 2007 s.1(4)(b).
[90] See Leigh, *Criminal Liability*; Coffee, 'No Soul to Damn, No Body to Kick', part II.
[91] Fisse and Braithwaite, 'Allocation of Responsibility'.

publication enhances openness. In assessing enforcement effectiveness, four main approaches can be used. These focus on: inputs; processes; outputs; and outcomes (See pages 35–6 above).

The practical challenges are significant, however. A study of Defra enforcement highlights a number of points.[92] First, accurate assessments of overall effectiveness in achieving outcomes cannot be made (even within a single-operator, single-tool regime) unless the regulator is able to calculate not only levels of non-compliance but levels of 'off-screen' non-compliance—errant behaviour which is beyond the reach of the regulatory regime, yet is relevant to the achievement of objectives. Second, clarity of legal and policy objectives is a precondition of effective assessment. Third, risk-based systems can provide a ready means of effecting year on year comparisons of performance—risk scores can be compared quite easily. Such systems, however, will not measure the effects of regulation on parties outside the system, and are quite easily manipulated by officials. Fourth, the natural inclination to focus on enforcement inputs (which offer cheaper, quicker, and more reliable statistics to be gathered) has to be balanced with efforts to measure outcomes on the ground. Fifth, in some regulated areas it is possible to identify 'short cut' or proxy measures of effectiveness—thus in relation to pesticides it may be feasible to analyse residues in water and use this as an indicator. Finally, where responsibilities for enforcement are unclear, or spread across numbers of institutions, this may impede the accurate assessment of effectiveness—because of coordination difficulties, institutional politics or divergencies in data collection and processing methods. Rationalization of regulatory responsibilities may accordingly offer ways to improve assessments, but only where, as noted, old coordination problems are not simply contained in a new organizational wrapper,[93] or rationalization does not produce its own.

Fisheries is an area that further illustrates the challenges of assessment. In fisheries regulation a key outcome measure is state of stocks, but this is affected by many factors other than enforcement.[94] Levels of compliance are also difficult to measure. As indicated, a considerable amount of non-compliant activity goes on beyond the inspection regime and the NAO reported in 2002–3 that it was impossible, in the then current system, to determine the number of undetected infringements.[95] These infringements related to compliance both with technical regulations and with the recording of landings. It was not physically possible to inspect enough vessels to ensure that landings were

[92] See R. Baldwin and J. Black, *Defra: A Review of Enforcement Measures and an Assessment of their Effectiveness in Terms of Risk and Outcome* (London, 2005).

[93] See, e.g., J. Black, 'Managing Regulatory Risks and Defining the Parameters of Blame: The Case of the Australian Prudential Regulation Authority' (2006) *Law and Policy* 1.

[94] e.g. global environmental changes—see Defra, *Review of Marine Fisheries*, 113.

[95] See NAO, *Fisheries Enforcement*, 2, 15–18.

accurately recorded.[96] Such difficulties drove the regulators towards secondary measures of effectiveness (e.g. probabilities of inspection)[97] and to data on processes and outputs (such as sea inspections, port visits, and prosecutions). As a result, Defra was ill-placed to measure the effectiveness of its detection system, its enforcement system, or its processes of assessment. Nor was it able accurately to judge the need to develop and apply new tools for detection, enforcement, or assessment.[98]

The lack of clear outcome objectives and benchmarks further undermined the assessment process in this area,[99] and a separate difficulty reported by the NAO was that EU Member States placed different interpretations on what constituted a serious infringement.[100] Even within English enforcement, infringements in different inspection districts were not recorded in a consistent manner. The NAO concluded that Defra was not able to monitor whether each district was dealing with infringements appropriately or to construct a picture of the nature or frequency of infringements so as to inform enforcement activity.[101]

Modification: The Adjustment of Tools and Strategies

The fifth core task within regulatory enforcement is, again, ongoing and involves modifying the regulatory approach in a manner that is informed by prior assessments of performance. Modification takes on board the adjustments of responses—the tools and rules that are used for both detection and compliance-seeking purposes and it also encompasses the modification of enforcement strategies themselves. As already suggested, modification also demands a willingness to think 'outside the envelope' and to consider whether, instead of adjusting the tools and enforcement strategies within the current regulatory strategy, it is necessary to effect a 'third order' or 'paradigm-shifting' change[102] by adopting a new regulatory (as opposed to enforcement)

[96] See NAO, *Fisheries Enforcement*, 3.

[97] Said by Defra to be 'probably the best readily obtainable measure of effectiveness'—*Review of Marine Fisheries*, 113.

[98] See R. Baldwin and J. Black, 'Defra: A Review of Enforcement Measures and an Assessment of their Effectiveness in Terms of Risk and Outcome' (London, 2005).

[99] See NAO, *Fisheries Enforcement*, 16; Prime Minister's Strategy Unit, *Net Benefits* (London, March 2004), 11.

[100] See NAO, *Fisheries Enforcement*, 16.

[101] See NAO, *Fisheries Enforcement*, 24.

[102] On the distinction between 'first-order changes' of regulation (e.g. tunings in the given regulatory control as exemplified by a change in the X in a price control formula) versus 'second-order changes' such as switches of instrument (e.g. from RPI-X to rate quantum price controls) versus 'third-order changes' or 'paradigm shifts' (e.g. abandoning command and control standards in favour

strategy (or mix of strategies)—for example, by moving from a state-imposed command and control centred regime to a completely different regulatory style such as one giving centrality to a scheme of industry-administered guidance and training.

Modification is an essential task, since there is only limited value in assessing performance if the regime is not to be adjusted so as to improve performance. Moreover, as the NAO report into fisheries also found, weaknesses in assessment systems can undermine capacities to modify processes.[103] In that sector, Defra was found by the NAO to operate inflexibly in its deployment of resources and people, which reduced its capacity to adjust its inspection strategies.[104] A special problem was lack of staff mobility which reduced operational responsiveness.[105] What was clear to the NAO was that a large number of strategic options were open to Defra but that these had not been fully assessed, explored, or put into effect.[106]

Conclusions

Enforcement tools are important, but the DREAM framework makes it clear that there are dangers in attempting to achieve better enforcement through a predominant emphasis on increasing the effectiveness of certain tools (e.g. criminal penalties). In many situations, the better way forward may be to improve detection techniques, response strategies (including approaches to selecting tools), performance assessment processes, and modification capacities. In the case of certain regimes, the most positive route to improved performance may involve thinking beyond enforcement of the current regime to broader issues of regulatory technique and institutional design.

Enforcement, can influence regulatory success or failure not merely by affecting the achievement of the right objectives. It can also impinge on the quality of regulatory processes. There is, however, as much art as science in enforcement since trade-offs have to be made on a number of fronts—between, for example, punishing infringers and maximizing compliance levels or between preventing creative compliance and producing rules that are easily enforced. In making these trade-offs, issues of accountability, due process,

of emissions trading) see J. Black, 'What is Regulatory Innovation?' in J. Black, M. Lodge, and M. Thatcher, *Regulatory Innovation* (Cheltenham, 2005).

[103] For discussion of changes post the NAO Report see Defra, *Review of Marine Fisheries*.

[104] See NAO, *Fisheries Enforcement*, 4.

[105] See NAO, *Fisheries Enforcement*, 35.

[106] But for subsequent action see Defra, *Review of Marine Fisheries*; Prime Minister's Strategy Unit, *Net Benefits*, and also Defra, *Securing the Benefits* (London, 2005).

and expertise arise. It may, for example, be necessary to use high levels of discretion in a regime of flexible and targeted enforcement if the 'right' results are to be produced, but questions of accountability and fairness are involved and it is proper that trade-offs with legitimate outcome gains should be argued out by regulators. Not only does enforcement demand that highly complex trade-offs and balances be carried out, it also demands that these be justified. The need for regulatory legitimacy, it should be emphasized, runs through the entire regulatory process.