

2004

## Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes

William H. Simon

*Columbia Law School*, [wsimon@law.columbia.edu](mailto:wsimon@law.columbia.edu)

Follow this and additional works at: [https://scholarship.law.columbia.edu/faculty\\_scholarship](https://scholarship.law.columbia.edu/faculty_scholarship)



Part of the [Jurisprudence Commons](#)

---

### Recommended Citation

William H. Simon, *Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes*, LAW AND NEW GOVERNANCE IN THE EU AND THE US, GRÁINNE DE BÚRCA & JOANNE SCOTT, EDS., HART PUBLISHING, 2006; COLUMBIA PUBLIC LAW RESEARCH PAPER NO. 04-79 (2004).

Available at: [https://scholarship.law.columbia.edu/faculty\\_scholarship/1341](https://scholarship.law.columbia.edu/faculty_scholarship/1341)

This Working Paper is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact [scholarshiparchive@law.columbia.edu](mailto:scholarshiparchive@law.columbia.edu).

# Columbia Law School

Public Law & Legal Theory Working Paper Group

Paper Number 04-79

## **TOYOTA JURISPRUDENCE: LEGAL THEORY AND ROLLING RULE REGIMES**

BY:

PROFESSOR WILLIAM SIMON  
COLUMBIA LAW SCHOOL

*This paper can be downloaded without charge from the  
Social Science Research Network electronic library at:*

*<http://ssrn.com/abstract=602626>*

## 2

*Toyota Jurisprudence: Legal Theory  
and Rolling Rule Regimes*

WILLIAM H SIMON\*

CONTEMPORARY ENGINEERING HAS produced distinctive ideas about organisation. These ideas have focused most on economic production, especially manufacturing, but they have broader relevance. I propose to show in this essay that a specific cluster of these ideas known as the Toyota Production System (TPS), or more generally, as 'lean production', contradicts basic premises of mainstream legal theory and implies a quite different jurisprudence.

I discuss TPS as if it were a legal system, but I am not interested in defending this characterisation on any but heuristic grounds. The jurisprudential implications of TPS are interesting because these ideas seem to underlie, in more fragmented and amorphous form, some important recent public law developments, especially in health-and-safety regulation and the delivery of public services. The engineering perspective thus has the potential both to clarify the basis of these developments and to suggest how they may evolve.<sup>1</sup>

The mainstream perspective to which I contrast Toyota jurisprudence is a set of premises common to the rights-and-principles and law-and-economics schools of legal theory, which together embrace most of the range of current American theory. As I use the term, however, mainstream jurisprudence does not include the tendencies some legal scholars call 'experimentalist', 'reflexive' or 'responsive' law; still others call 'new governance', and had the term not been lately abused, might usefully be called pragmatist.<sup>2</sup> The Toyota

\* Arthur Levitt Professor of Law, Columbia University. Thanks for help and encouragement to Paul Adler, Barbara Fried, the Hewlett Foundation, Chuck Sabel, Joanne Scott, Neil Walker, and participants in discussions at Cambridge University, Columbia Law School and the State University of New York at Buffalo Law School.

<sup>1</sup> The idea that TPS and related engineering ideas have important implications for public institutions comes from Charles F Sabel, 'Learning by Monitoring: The Institutions of Economic Development' in Neil Smelser and Richard Swedberg (eds) *Handbook of Economic Sociology* (Princeton, NJ: Princeton University Press, 1994).

<sup>2</sup> See eg, Michael C Dorf and Charles F Sabel, 'A Constitution of Democratic Experimentalism' (1998) 98 *Columbia Law Review* 267; John Braithwaite, *Restorative Justice and Responsive Regulation* (Oxford: Oxford University Press, 2002); Joanne Scott and David M Trubek, 'Mind the Gap: Law and New Approaches to Governance in the European Union' (2002) 8 *European Law Journal*. 1. The case for the term pragmatism rests on the convergence

38 *William H Simon*

perspective resonates with these latter tendencies. It has in some cases influenced work with these tendencies, and it may prove useful in elaborating them.

In contrast to mainstream legal thought, the Toyota system (1) emphasises the goals of learning and innovation (rather than of dispute resolution and the vindication of accepted norms); (2) combines the normative explicitness associated with formal rules with the continuous adjustment to particularity associated with informal norms (no dialectic of rules and standards); (3) treats normative decision making in hard cases as presumptively collective and interdisciplinary (rather than the heroic labour of a solitary professional); (4) fosters a style of reasoning that is intentionally destabilising of settled practices (rather than harmonising or optimising); and (5) attempts to bracket or sublimate issues of individual and retrospective fairness or blame.

Section 1 of this chapter is a stylised account of a cluster of basic themes in mainstream jurisprudence. Section 2 describes the contrasting themes of Toyota jurisprudence. Section 3 shows that the Toyota themes are visible in recent American public law developments, especially in health and safety regulation and school reform. It also discusses proposed reforms of death penalty administration in order to rebut the intuition that the Toyota perspective would be out of place in areas that are strongly charged morally. Section 4 speculates briefly about the general direction that legal theory would take if it absorbed the insights of the Toyota perspective.

## 1. MAINSTREAM JURISPRUDENCE

Most legal thought is committed to the following five propositions:

*A. The basic functions of the legal system are dispute resolution and the vindication or optimisation of accepted social values.*

In the rights-and-principles perspective, the legal system is, in HLA Hart's formulation, a 'union of primary and secondary' norms.<sup>3</sup> The primary norms prescribe rights, duties and powers for citizens in civil society. The efficacy and legitimacy of these primary norms depends on their convergence with informal social norms. Secondary norms allocate authority among officials to resolve disagreement and conflict about what the primary norms should be and how they should be applied and enforced. No matter how perfect the convergence between primary rules and informal social norms, disputes will arise because of bad faith or the incompleteness of social morality, and secondary norms are needed to resolve such disputes. The secondary norms may themselves express social values such as due process or democracy, or they may rest simply on the social interest in the

of some of the initiatives discussed here with John Dewey's conception of democratic government as a process of participatory learning. See John Dewey, *The Public and Its Problems* (Athens, OH: Swallow Press, 1954 [1927]).

<sup>3</sup> HLA Hart, *The Concept of Law* (Oxford: Clarendon Press, 1961) 77, 92–93.

resolution of conflict. Lawyers who see a broad and rich social consensus, or who cannot imagine a legitimate legal order without one, tend to see primary norms as fundamental. Those who focus on the fact of conflict and disagreement about primary norms tend to emphasise secondary norms.<sup>4</sup>

In the law-and-economics perspective, the distinction between primary and secondary norms is less central, and norms tend to merge with a broader array of preferences, but the picture is similar. The legal system is supposed to maximise welfare, understood in terms of current preferences. The system is an instrument for satisfying these preferences. So the task is to trade off the benefits of increased preference satisfaction through law making and enforcement against the costs of these activities so as to attain the greatest net benefit. Conflict and disagreement are treated as either costs of enforcement or as pre-existing social harms, in either case as something to be minimised.<sup>5</sup>

*B. The dialectic of rules and standards: The legal system chooses formal rules in order to limit discretion but at the price of reducing sensitivity to factual particularity; it chooses informal norms (standards) to make decisions sensitive to context but at the cost of reduced control of discretion.* Formal norms constrain decision more than informal ones. They limit the range of factors that the decision maker can consider. Legal theory attributes several potential advantages to formality, but the advantage most emphasised in general legal discourse is that formality inhibits the decision maker's ability to make decisions on the basis of illegitimate considerations—her personal or political values and goals. The narrower and more explicit the rule, the less room it allows for influence by personal values and goals, and the more readily a reviewing authority can determine whether the decision conforms to the rule.<sup>6</sup>

But of course, the more the rule constrains discretion, the more it also inhibits consideration of how its purposes will apply in the full context of the decision and the more often it will produce decisions at variance with these purposes. The gap between the norm's underlying goals and the decisions it produces is directly proportional to the norm's formality.

Lawyers generally tend to favor discretion-constraining rules for groups they distrust and discretion-conferring standards for groups they trust, though they do not always agree on the identity of the groups. During the Warren Court era, liberal lawyers favoured rules for the street-level public workforce—for example, *Miranda* rules for the police—and standards like 'just cause' for judges. More recently, conservatives who distrust the judiciary in criminal matters have sought to restrict their discretion through the

<sup>4</sup> Compare Hart previous n (emphasising secondary norms) with Ronald Dworkin, *Law's Empire* (Cambridge, MA.: Harvard, 1986) 176–275 (emphasising primary norms).

<sup>5</sup> Eg, Richard Posner, *Economic Analysis of Law*, 6th edn (New York: Aspen, 2003) 215–27, 563–65, 595–98.

<sup>6</sup> See generally, Frederick Shauer, *Playing by the Rules: A Philosophical Examination of Rule-based Decision-making in Law and Life* (Oxford: Oxford University Press, 1991) 149–55; Posner, previous n, at 556–57.

40 *William H Simon*

rules of the Federal Sentencing Guidelines, with the tacit but predictable effect of increasing the discretion of prosecutors, a group they trust, whose charging decisions are regulated by standards.

All these lawyers, however, tend to view the choice between rules and standards as primarily a trade-off between the costs and benefits of limiting discretion on the one hand and of decisional contextualisation on the other.

*C. Legal reasoning in hard cases is interstitial.*

It is a matter of pride in our legal tradition that the courts exercise their law-making power ‘one case at a time’.<sup>7</sup> They tailor their decisions to the facts of the dispute before them. The decisions are conclusive only with respect to those facts. Their precedential effect for other cases takes the form of ‘gravitational force’ proportional to the proximity of their facts to the ones in the decided case.<sup>8</sup>

Hard cases are those for which no statute or past case plainly dictates an answer. We resolve hard cases by deriving results that harmonise with surrounding authority. We accept as presumptively valid the clear answers the authority gives for other cases and try to derive principles from them that are consistent with each and indicate how the case at hand should be resolved. A good answer is one that fits with the surrounding authority.

Ronald Dworkin’s rights-and-principles account portrays ‘fit’ or ‘integrity’ as a virtue in itself. However, there are two other rationales for interstitial decision making that are widely accepted within both the rights-and-principles and law-and-economics perspective. The first emphasises the limited capacities and democratic legitimacy of courts. We need courts to resolve disputes, but their ability and authority to promulgate rules that go beyond the specifics of these cases is controversial. The second emphasises the importance of predictability. Both fairness and efficiency are served when citizens can predict when the state will intervene in private affairs. Such predictability allows successful planning of individual activity and successful coordination of joint activity. In this view, the law strives to eliminate the friction of individual interaction. This goal is served when courts respect settled precedent and innovate only when necessary (either because the law ‘runs out’ or because social change has rendered settled precedent intolerable).<sup>9</sup>

*D. The core form of legal decision making is strongly differentiated from other roles and activities.*

Legal theory has been most preoccupied with explaining a single phenomenon—how the judge pulls the rabbit of determinate resolution out of the hat of ambiguous authority. It treats other types of legal work as peripheral,

<sup>7</sup> Eg, Cass Sunstein, *Once Case at a Time: Judicial Minimalism on the Supreme Court* (Cambridge, MA: Harvard, 2001).

<sup>8</sup> Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard, 1978) 111; see also Dworkin, n 4 above, at 250–54 (on the principle of ‘local priority’ in the assessment of precedent).

<sup>9</sup> Posner, n 5 above, at 560–61.

and it treats the judge's work as distinct from and independent of that of other actors even within the court system.

The rights-and-principles tradition argues that judicial decision making is strongly differentiated methodologically from other forms of public decision making. To a unique extent, judges must reason prescriptively from authority, rather than instrumentally with reference to goals. And the judge's isolation from other officials and her relative lack of accountability for the practical effects of her decisions are sometimes praised as a key elements of the judicial 'independence' that is a key feature of the rule of law.

For other theorists, the tendency to treat the judge's role in institutional isolation is more a tacit, or even regretted, habit than a matter of principle, but the habit is pervasive and longstanding. Legal theory takes little account, even within the legal system, of anyone but judges, even though we know that the practical effects of judicial intervention depend on clerks, probation officers, special masters, bailiffs, sheriffs, marshals, trustees, executors and receivers. The theorists' interest in the practices of courts typically stops at the point of judicial decision. Whether and how the decision is enforced or its broader practical impact is a matter of secondary, if any, concern. Judicial decisions are often appraised in terms of their internal plausibility, rather than in terms of their practical contribution to realising their goals.<sup>10</sup>

It is often pointed out that this focus on the judicial decision is parochial and anachronistic. Even within the court system, judicial activity is increasingly described as 'managerial', especially in connection with two kinds of cases. First, there are institutional reform cases in which the courts undertake to reform the administrative structures and practices of public institutions like schools and prisons that have failed to discharge their duties. Second, there are cases like those involving small drug crimes or child welfare—small cases that occur in large numbers with common patterns in which the court is expected to intervene in a fairly complex long-term manner in the hope of improving the welfare of some of the parties. Although these developments are widely recognised, they have not much affected the preoccupations of legal theory.<sup>11</sup>

*E. Legal decision making is paradigmatically retrospective and individualist.* Fairness is a central preoccupation of American common law decision making. Fairness means basically corrective justice. It is centrally concerned with

<sup>10</sup> There is, of course, a substantial empirical and instrumental literature on the effects of legal rules and judicial decisions. Much of this literature is theoretical and might well be called 'legal theory'. However, I exclude it from the definition of 'mainstream legal theory' because it is more peripheral to the legal culture than the judge-focused rights-and-principles and law-and-economics literature and has had limited influence on them.

<sup>11</sup> See Judith Resnik, 'Managerial Judges' (1982) 96 *Harvard Law Review* 374; Michael C Dorf and Charles F Sabel, 'Drug Courts and Emergent Experimentalist Government' (2000) 53 *Vanderbilt Law Review* 831.

42 *William H Simon*

the allocation of benefits and burdens among individuals, and with linking present treatment to past conduct. The American legal system stands ready to commit vast resources to the determination and evaluation of past conduct in order to calibrate present reward or punishment to it. The concern with fairness is longstanding, but it has recently intensified because of the increased punitiveness of the legal system in both its civil and criminal spheres. Punishment is usually predicated at least in part on retrospective fairness values, and it triggers demands for more elaborate efforts to make the relevant determinations.

Fairness is the fundamental normative commitment of the rights-and-principles school. These theorists are prone to defend fairness values as more fundamental than others or to suggest that they are more strongly suited to vindication by judges.<sup>12</sup> Law-and-economics is different on this point. It rejects fairness in favour of welfare as a normative touchstone. Yet, despite this distinction, law-and-economics also focuses on crafting legal rules that apportion current reward and sanctions to past individual conduct.<sup>13</sup> Sometimes fairness values re-enter in the form of preferences, which then become part of the welfare function to be optimised. Sometimes the results dictated by corrective justice norms turn out to be the best means of regulating behaviour in ways desirable on welfare grounds. Even when its prescriptions depart from the rights-and-principles view, law-and-economics most often proceeds from an individualist and retrospective approach. It seeks to craft incentives that induce individual, presumptively rational, actors to behave efficiently. And it accepts the common law premise that future behaviour is effectively regulated through the design of sanctions to fit past conduct in current cases.

Two problems in contemporary legal systems challenge the idea of retrospective individual liability in either the direct fairness or efficiency perspectives. The first is that much conflict arises as a by-product of conduct that is generally socially valuable. The paradigmatic traditional cases of individual retrospective liability involve intentional instances of generally noxious behaviour, say, battery or theft. But a very large class of contemporary cases—including unintentional torts and many violations of business regulation—involve conduct that is in general beneficial but that has been carried too far or conducted with insufficient precautions, often inadvertently. Legal theory has wrestled with the question of when such conduct should be regarded as blameworthy or inefficient. A key touchstone has been the idea of negligence or reasonableness. The meaning of the reasonableness norm is contested. One ambiguity refers to the relation between customary practice on the one hand and optimal practice on the other. Sometimes

<sup>12</sup> Dworkin, n 8 above.

<sup>13</sup> Louis Kaplow and Steven Shavell, *Fairness versus Efficiency* (Cambridge, MA: Harvard, 2002) 134–48.

reasonableness means the dominant practice among similar actors. Sometimes it means something more. But in either case, the inquiry in mainstream legal thought remains retrospective and individual. It is the prevalent or optimal practice at the time of the episode that counts, and whatever standard we come up with must have been within the capacity of the defendant at the time.<sup>14</sup>

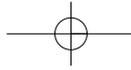
Yet, it is often debatable whether either fairness or efficiency warrants liability based on such a retrospective inquiry. To many, fairness values do not warrant sanctions for a broad range of harms arising from unforeseen consequences of generally beneficial conduct that are nevertheless swept into our liability system. Others suggest that imposing liability may be inefficient because material compensation does not effectively redress symbolic or emotional injuries, or because the liability process does not directly induce improvements in the systemic practices that produce harm, and, to the extent that it causes actors to hide information, may inhibit improvement.<sup>15</sup>

A second problem arises from the fact that the defendants in a large fraction of legal controversies are organisations. It is often not clear how either fairness or efficiency dictates the imposition of liability on organisations. One strategy is to ignore or pierce the organisation and impose liability on its individual constituents when their conduct would independently warrant it. Another is to treat the organisation as if it were an individual when the elements of liability can be established on the basis of the aggregate conduct of multiple constituents acting for the organisation. Neither strategy is entirely satisfactory. With the first, decision making remains plausibly individualistic but at the cost of ignoring the distinctive dangers and problems of collective activities. With the second, the individualism seems strained. For example, while the criminal punishment of corporations does not strike most people as per se unjust, it is very hard to explain in terms of individualist fairness notions when, for example, the main effect falls on current employees and shareholders who neither encouraged nor benefited from the wrongdoing.

Moreover, while law-and-economics scholarship often treats the incentive effects of liability on organisations as tantamount to the effects on individuals, this is usually a matter of faith more than logic or evidence. Economics presupposes individual decision makers, and corporations are complex aggregations of individuals. Nevertheless, the practice of either ignoring organisations or pretending that they are individuals has long been favoured by the legal culture as a way of reconciling its individualistic doctrines with the reality of a world of organisations.

<sup>14</sup> 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)) (Mineola, NY: Foundation Press, 1994) 403–49.

<sup>15</sup> See, eg, Braithwaite, n 2 above.

44 *William H Simon*

## 2. THE TOYOTA PRODUCTION SYSTEM AS A JURISPRUDENTIAL PHENOMENON

TPS arises from dissatisfaction with a traditional mass manufacturing model that combines central planning with ad hoc shop floor adjustment.<sup>16</sup> In the traditional model, a central corps of managers and engineers promulgates rules that dictate practice to a workforce that is narrowly skilled and divided among functional departments (for example, milling, painting, cutting, polishing, assembly). Central management forecasts sales and then prescribes production targets for specific products, orders materials, and schedules each phase of the production process. Typically, the plan calls for the parts of a product to be processed separately in different departments in large batches with specialised machinery.

Invariably, adjustments are required as events depart from the plan. The pattern of orders is different from the forecast. Supplies fail to arrive on time, or they are defective. In the plant, machines break down, or parts are improperly machined. The plan contemplates a series of adjustment mechanisms for such contingencies. 'Expeditors' may travel between the sales department and the factory floor to advance orders ahead in the production queue in accordance with the needs and clout of the customer. The plant can maintain a central parts inventory, or each worker can have a 'buffer' stock of spare parts at her station in case new parts fail to arrive on time or in proper condition. A specialised inspectorate or Quality Control department can review finished product and send non-conforming items back or to a special department for re-work.

TPS proponents complain that such a system is slow in responding to unanticipated changes in the volume of customer demand or in its capacity to modify or change products. It takes a long time for centralised management to absorb information indicating that changes are needed, and a long time for it to develop and implement needed changes. The traditional system tends to be quite wasteful of labour and materials, in part, because it is slow to discover defects and, when they are discovered, slow to remedy them. In addition, the system does not effectively capitalise on the knowledge and potential creative effort of most its workers. And by encouraging tolerance for errors and the expectation that they will be remedied downstream, the system fails to cultivate in workers a sense of responsibility or 'ownership'.

TPS tries to reduce reliance on forecasting, first, by configuring its sales efforts to smooth out orders, and second, by making the manufacturing process more flexible and quick. Production cues are transmitted

<sup>16</sup> For general accounts, see Taiichi Ohno, *Toyota Production System: Beyond Large Scale Production* (Portland, OR: Productivity Press, 1988); Shigeo Shingo, *Study of Toyota Production System from an Industrial Engineering Point of View* (Portland, OR: Productivity Press, 1981); James P Womack, Daniel T Jones, and Daniel Roos, *The Machine that Changed the World: The Story of Lean Production* (New York: Macmillan, 1990).



less hierarchically through central planners and more laterally through the shop floor. This is accomplished in part by the *kanban* system in which downstream stations signal their need for parts by returning cards upstream. Parts and products tend to be produced in smaller lots in a more continuous process with more broadly skilled workers and less specialised machinery.

The traditional model strongly differentiates between conception, which is the responsibility of central management, and execution on the shop floor. TPS attenuates this distinction. The phrase *kaizen*, or continuous improvement, connotes that process be revised in the course of its execution. TPS diffuses responsibility for the organisation of production broadly. Shop floor teams write and revise the descriptions for their jobs, schedule their members, and arrange for maintenance and repair of their equipment. Some workers meet regularly in 'quality circles' to consider problems and propose solutions. All workers are encouraged to make suggestions for improvements in the process, and such suggestions often result in changes. Inspection and quality control cease to be the exclusive preoccupation of an elite corps and become the responsibility of the entire workforce. While the old system tolerated a significant number of defects and expected many items to be re-worked, TPS espouses 'zero tolerance' for defects and has no re-work department.

The *kaizen* idea and the 'zero tolerance' norm dictate a distinctive response to production problems. A problem occurs when a defective part appears, or when a worker is unable to perform a prescribed action (say, because there is insufficient time), or when performing the prescribed action would be inconsistent with quality norms (for example, painting over dirt). The system discourages ad hoc adjustments to such problems. It does not permit workers to pass defective parts down the line, and it does not provide buffer inventory to substitute for defective parts. Instead, such problems are treated as symptoms of structural flaws to be remedied immediately. If necessary, this means stopping the line entirely until the problem is fixed. Shop floor workers have the power and duty to pull the '*andon* cord' that brings production to a halt and illuminates an electronic display that signals to the entire plant that a problem has occurred. In principle, production resumes only when the process has been revised sufficiently to eliminate the danger of recurrence. The procedure through which the solution is devised is one of consensual deliberation by representatives of any group likely to have relevant knowledge or whose cooperation is likely to be necessary to implementing the solution.

If we consider TPS as a legal system, the following points emerge:

*A. Learning and innovation are key purposes of the system.*

The purposes of TPS are not adequately accounted for by the preoccupations of mainstream legal theory. TPS has primary norms that reflect a pre-existing consensus about legitimate behaviour—rules that prohibit assault and theft, for example. It also has secondary norms that constitute hierarchical procedures of coercive dispute resolution. Neither kind of rule, however,

46 *William H Simon*

plays a salient role in a successful plant, and neither is directly associated with what the participants consider the most important and distinctive functions of TPS.

From the participants' point of view, the key norms are those that constitute the problem-solving process. This process is not concerned with dispute resolution. It starts out with a shared sense among the participants that something is wrong and a shared sense of common goals. What requires discussion is not disagreement, but uncertainty. Moreover, neither the composition nor the operation of the problem-solving groups is determined by norms about authority.<sup>17</sup> Composition is determined by the principle that 'anyone with relevant knowledge of a problem is included, regardless of rank'.<sup>18</sup> Process is a matter of consensus, not a search for a pre-existing consensus, but an effort to forge a new one.

The problem-solving discussions do take place against a background of general values of 'quality'. But it would be misleading to see the discussions as efforts to vindicate or maximise these values. The values are indeterminate. The discussions are as much efforts to define them as to implement them.

For example, although early TPS proponents sometimes equated 'quality' with cost minimisation, the more recent literature speaks with contempt of 'cost-based strategies'.<sup>19</sup> The change apparently reflects the experience that American managers applying cost minimisation rigorously have tended to reject Toyota-style practices. They found that the Toyota approach tended to raise the costs that can most readily be measured—short-term direct labour costs—and to create benefits that are speculative (savings from future innovation) or hard to measure in advance (enhanced customer satisfaction and market share).

Similarly, 'quality' strategies like TPS are considered antithetical to managerial strategies preoccupied with stock price maximisation. Whether or not stock price maximisation is plausible goal, the strategy suffers from indeterminacy; the goal does not translate clearly into more concrete directives. In practice, it is associated with a focus on financial indicators, but

<sup>17</sup>

While lower levels [in TPS] did not have much authority to make decisions without prior consultation with superiors, this apparent centralization usually took the form of 'fact-based' dialogue based on expertise rather than command-and-control domination based on positional authority...Japanese organizations had more *de jure* centralization but also more *de facto* participation than comparable American organizations.

Paul Adler, 'Flexibility versus Efficiency: A Case Study of Model Changeovers in the Toyota Production System' (1999) 10 *Organization Science* 43 at 54.

<sup>18</sup> John Paul MacDuffie, 'The Road to Root Cause: Problem-Solving at Three Auto-Assembly Plants' (1997) 43 *Management Science* 479 at 495. My description of TPS is based largely on the espoused principles of its proponents. Studies such as MacDuffie's and Adler's, previous n, indicate that some plants conform extensively to these principles but also that there is a good deal of variation in practice among plants that purport to have absorbed them.

<sup>19</sup> Cf. Ohno, n 16 above, at 8–9, 52–56 (discussing cost minimisation as fundamental goal of the system) with Rajan Suri, *Quick Response Manufacturing* (Portland, OR: Productivity Press, 1998) 54–56, 76–78 (disparaging 'cost-based organization').

the stock price performance of the firms that have taken this tack has not been superior. Paradoxically, it appears that the managers most devoted to pleasing the stock market are not the ones who end up doing so, at least in the long run.<sup>20</sup>

A more popular definition of quality associates it with customer satisfaction. But what products and product features customers value and how customers trade off performance and price are matters that are not fully known in advance. They are discovered in the process itself.

The quality norm is thus not a consensus to be vindicated or a value to be maximised. It is both a bet and a cultural commitment: a bet that short-term process investments will lead to discoveries that will shift cost curves; a cultural commitment that unifies and motivates the disparate groups in the production process by giving creativity and social dignity to their work.

The dimension of TPS omitted in mainstream jurisprudence is the focus on learning and innovation. TPS is, most distinctively, an effort to discover new ways of producing. The diffusion of responsibility and the stressing of the system to induce constant collaborative revision force pooling of knowledge in ways intended to generate new understanding.

*B. TPS combines a strong commitment to formalisation of norms with continuous adjustment to unanticipated particularity.*

TPS norms are always as articulated as possible, but they are not applied consciously in a way that would frustrate their purposes. Instead, the rules get re-considered and re-written when they come in tension with unanticipated contingency. The workers' power/duty to pull the *andon* cord and stop production represents the authority to trigger a legislative deliberation on the amendment of the rule.

The purpose of formality in TPS is quite different from its primary purpose in mainstream legal thought. Toyota formality is not designed to restrict discretion. TPS is a high-trust system with broadly educated workers who have a strong general stake in their jobs and are subject to constant peer pressure to perform well. Rather, the purpose of formal norms is to facilitate learning, or as Michel Greif puts it, 'to inspire improvements'.<sup>21</sup>

'Say what you do, and do what you say' is a basic premise of Toyota-style engineering.<sup>22</sup> As it rejects the command-and-control model of Fordist bureaucracy, TPS rejects the traditional artisanal vision in which work is

<sup>20</sup> James Belohlav, 'Quality, Strategy, and Competitiveness' (1993) 32 *California Management Review* 55 at 56–61. See also Masaahi Imai, *Kaizen: The Key to Japan's Competitive Success* (New York, McGraw-Hill, 1986) 49–50:

Japanese managers have found that seeking improvement for improvement's sake is the surest way to strengthen the company's overall competitiveness. If you take care of quality, the profits will take care of themselves.

<sup>21</sup> Michel Greif, *The Visual Factory: Building Participation Through Shared Information* 74 (Portland, OR: Productivity Press, 1991).

<sup>22</sup> Joseph Casio, (ed) *The ISO 14000 Handbook* (Milwaukee, Wisconsin: American Society for Quality, 1996) 196.

48 *William H Simon*

regulated by tacit norms that can be grasped only by prolonged socialisation into guild and local workplace cultures. The duty to articulate forces the actors to reflect on what they are proposing to do and to communicate it as precisely as possible to their peers. Codifying the practice means that it can be more readily learned by people who did not participate in its formulation, or who do not share background experiences with the authors. The more explicitly defined the job, the broader the range of workers who can perform it. Thus it becomes easier to redeploy workers in accordance with their availability. It also increases the range of perspectives that can be brought to bear on the norm's revision. An articulated norm is more readily criticised and debated both because and it is more easily grasped and because criticism is less likely to risk personal offence when the norm is divorced from particular people.<sup>23</sup>

Formalisation is thus strongly associated with revisability. In TPS, a rule is simultaneously 'a point to adhere to and a point of departure'.<sup>24</sup>

*C. Decision making is fundamentally collaborative and interdisciplinary.* TPS, as we have seen, precludes the individual discretionary decision making involved in ad hoc adjustment to unanticipated particularity. On the other hand, it depends heavily on a different kind of judgment to revise continuously its rules and practices. This judgment is typically collective and interdisciplinary.

In traditional manufacturing, the process is often organised functionally, with each group or department focused on a particular activity, such as painting or machining or assembly. In TPS, production is organised more in terms of types of products, with each group performing a series of functions, and perhaps producing entire products, with a range of skills. Work is conducted by teams whose members are broadly trained and who make decisions collectively about job design, scheduling, and materials requirements. In traditional manufacturing, major unexpected problems are often handled exclusively by corps of specialists trained in a particular discipline, such as electricians or machinists. In TPS, when major problems arise, they are addressed by groups broadly representative of affected constituencies throughout the plant. These may include various shop floor teams, supervisors, suppliers, marketing people and designers.

Collaborative decision making responds to two constraints. First, the dispersal of information throughout the production process. No individual

<sup>23</sup> See Imai, n 20 above, at 74–78; Greif, n 21 above, at 74–76.

<sup>24</sup> *Ibid* at 75. See also, Paul Adler, 'Time and Motion Regained' (1993) *Harvard Business Review* 97 at 104,

[W]orkers create a consensual standard that they teach to the system by writing job descriptions. The system then teaches these standards back to workers, who, then, by further analysis, consultation, and consensus, make additional improvements. Continual reiteration of this disciplined process of analysis, standardization, re-analysis, refinement, and restandardization creates an intensely structured system of continuous improvement.

knows enough to make key problem-solving decisions and no fixed management group could be relied on to know enough to solve all problems. Second, the need for multiple perspectives. Many problems call for multiple kinds of technical knowledge or for kinds of technical knowledge that cannot be identified at the outset. Dirt on a component might be a machining issue (if it can be eliminated in the course of machining) or a painting issue (if it can be painted over so as to make it unnoticeable) or a marketing issue (if the customer will not care about it). Team decision making increases the range of both information and technical perspectives that are considered in making the decision.

*D. Decision making in TPS is intentionally destabilising.*

John Womack and Daniel Jones contrast TPS with 'steady state' management.<sup>25</sup> In the latter, the goal is to minimise disruption and, when disruption is unavoidable, restore equilibrium as quickly as possible. By contrast, TPS involves deliberate destabilisation.<sup>26</sup> It induces problems with highly specified rules, unforgiving quality standards, minute synchronisation, and minimisation of inventory.

Two practices associated with TPS are relevant here. The first is benchmarking—self-assessment in terms of the performances of industry leaders. Benchmarking subjects every practice and product to comparison with the firm's most successful competitors. This approach represents a shift from the traditional practice of comparing performance to industry averages—the in-house analogue to the common law reasonableness norm. Reasonableness connotes that typical performance should be maintained in steady state. But benchmarking connotes that only superior performance is adequate. Anything less requires efforts to improve, and in a competitive market, even superior performance is unlikely to be maintained without efforts to improve. Thus, the lean production literature is full of statements disparaging adherence to custom or reasonableness. For example, Womack and Jones write:

[T]he high achievers set specific timetables to accomplish seemingly impossible tasks and then routinely meet or exceed them. The low achievers, by contrast, ask what would be reasonable for their current organization ... to accomplish, and generally defeat themselves before they set out.<sup>27</sup>

<sup>25</sup> James P Womack and Daniel T Jones, *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, rev edn (New York: Free Press, 2003).

<sup>26</sup> This is not to say that TPS plants are less stable, merely that an important degree of instability is deliberate. Disruption is typically experienced as frequent when TPS methods are inaugurated but declines when they are successful:

Today, in Toyota plants, where every worker can stop the line, yields approach 100 percent. That is, the line practically never stops! (In mass-production plants by contrast, where no one but the line manager can stop the line, the line stops constantly ... to deal with material supply and coordination problems.) Womack, Jones and Roos, n16 above, at 57.

<sup>27</sup> Womack and Jones, n 25 above, at 95.

50 *William H Simon*

As a descriptive matter, such rhetoric needs to be discounted for hyperbole. However, the rejection of the goal of equilibrium and the commitment to continuous destabilisation it implies is a critical characteristic of the TPS strategy.

The other practice associated with destabilisation is ‘root cause analysis’. Problems are supposed to be traced backward through the production process. The rule-of-thumb prescribes tracing back five stages—the ‘Five Whys’.<sup>28</sup> This means that both inquiry and solution potentially ramify beyond the point at which the problem arose. For example:

Why is machine A broken? Because no preventive maintenance was performed.

Why was the maintenance crew derelict? Because it is always repairing machine B.

Why is machine B always broken? Because the part it machines always jams.

Why does the jam recur? Because the part is warped by heat stress.

Why does the part overheat? A design flaw.<sup>29</sup>

The point is that the solution to the problem may require unsettling practice at a point remote from the one at which it was identified. In the short term, this approach is often more disruptive and costly than narrower responses. John Paul MacDuffie illustrates this in a comparative study of American auto plants. He treats as evidence of scant progress in implementing lean production at a General Motors plant that a quality auditor’s response to a paint defect was limited to ‘charg[ing] the costs’ of the defect to an earlier inspection team that should have discovered it.<sup>30</sup> This limited response violates the spirit of TPS. It does not inquire into the remote cause of the problem, or even of the inspectors’ failure. It relies on indirect inducements to improve performance—penalties to inspectors—that, if they have any effect at all, are likely to cause low-visibility ad hoc adjustments, rather than explicit systemic revision.

The problem-solving process associated with root cause analysis is incremental in the sense that it begins with dissatisfaction with only a small part of the process and ends with a revision of only a small part. But it is not interstitial in the common law sense of ‘one case at a time’. With root cause analysis, the initial definition of the problem does not control its analysis or disposition. Every problem potentially implicates any of the parts of the system.

<sup>28</sup> Ohno, n 16 above, at 17–18.

<sup>29</sup> The phrase ‘root cause analysis’ is a little misleading to the extent that it implies that we are interested in a historical or scientific account of how the error happened. The search is, not so much for a retrospective description, as for opportunities for improvement. The ‘five’, of course, is a crude rule of thumb. Its blatant arbitrariness is an acknowledgment that the search for cause is potentially infinite.

<sup>30</sup> MacDuffie, n 18 above, at 486.

This insistence that intelligent implementation of the parts of the system requires an understanding of how each part fits into the whole is reflected in the broad training workers receive and in plant design for transparency. Plants are laid out simply and logically. Clutter and visual obstructions are minimised. Elaborate displays visible from all points in the plant summarise what is happening at each station. When a problem that requires suspension of production occurs, its nature and location are communicated immediately to the entire plant. The premise is that, at the outset, we cannot say which people in the plant will have the knowledge and skills necessary to the solution.<sup>31</sup>

The Toyota perspective thus resonates with a point Dworkin makes against Cass Sunstein's defense of the 'one case at a time' approach to adjudication.<sup>32</sup> Sunstein argues that judges do better to resolve cases in terms of relatively narrow, low-level principles. Relatively high-level principles should be disfavoured because they have implications that extend broadly across the legal system, and thus require more complex decision making, have greater potential to produce disagreement, and are more likely to upset settled expectations. Dworkin replies that we expect judges to justify their decisions in terms of principle, and there is no reason to think that the principles that provide the best justifications for decisions will be consistently of limited generality. What Dworkin calls 'justificatory ascent' means that particular claims of right potentially implicate broader structures in a way analogous to the ramifying tendency of Toyota production problems.

<sup>31</sup> The Toyota perspective challenges divisions, not only between work roles and groups within the firm, but also the boundaries between firms. The challenge of continuous improvement is applied to the entire process of producing a given product or product family. Just-in-time delivery of components requires close coordination with suppliers. The upstream 'root cause' of a problem may often lie beyond the boundaries of the firm. The current inter-firm divisions are provisional, and since firm boundaries tend to obscure information, suspect. Thus, the Toyota approach prescribes close collaboration and free information-sharing between firms at different stages of the production process. It urges process designers to envision and perfect the entire 'value stream' without regard to how it is currently partitioned:

For example, when Pratt & Whitney, the world's largest manufacturer of aircraft jet engines, recently started to map its value streams for its families of jet engines, it discovered that activities undertaken by its raw materials suppliers to produce ultrapure metals were duplicated at great cost by the next firms down stream, the forgers who converted metal ingots into near-net shapes suitable for machining. At the same time, the initial ingot of material—for example, titanium or nickel—was ten times the weight of the machined parts eventually fashioned from it. Ninety percent of the very expensive metals were being scrapped because the initial ingot was poured in such a massive size—the melters were certain that this was efficient—without much attention to the shape of the finished parts. And finally, the melters were preparing several different ingots—at great cost—in order to meet Pratt's precise technical requirements for each engine, which varied marginally from those of other engine families and from the needs of competitors. Many of these activities could be eliminated almost immediately with dramatic cost savings. (Womack and Jones, n 25 above, at 20).

<sup>32</sup> Ronald Dworkin, 'In Praise of Theory' (1997) 29 *Arizona State Law Journal* 353 at 368–75.

52 *William H Simon*

But the analogy ends when Dworkin describes the goal of decision making as establishing harmony within an existing body of norms. In TPS, legitimacy does not depend on connecting decisions to established authority. (To be sure, there will usually be some such connection, but not because legitimacy requires it; rather because, as Stanley Fish emphasises, the decision makers cannot escape the influence and constraints of their cultural and practical circumstances.)<sup>33</sup> In TPS, legitimacy arises from the provisional consensus of the stakeholders and their commitment to re-assess readily in accordance with agreed-on criteria.

Another destabilising aspect of Toyota-style decision making arises from the interaction of its prescriptive and instrumental aspects. Rights-and-principles reasoning is prescriptive; it is concerned with elaborating goals—values with intrinsic worth. From its perspective, implementation is secondary. Law-and-economics reasoning is instrumental; it focuses on means and implementation. It usually treats the goals it seeks to further as given and fixed. In contrast, reasoning in the Toyota system encourages the re-assessment of ends in the process of implementing them. Discussion starts out as instrumental to, inter alia, defined goals of product quality, but these goals can be reconsidered in the course of problem-solving. In traditional plants that have experimented with shop floor teams and quality circles, discussion stops when it reaches ‘design issues’<sup>34</sup> But in plants that have gone farther, design engineers may be dispatched from headquarters and assigned to work with teams in the plants. Thus, the difference between production (means) and design (goals) is eroded. Decision making in manufacturing is potentially destabilising, not only of spatially remote features of the production process, but of settled expectations about the goals of the process.

The interaction of prescriptive and instrumental thought is also salient when new products are being designed. The traditional process is a sequential and departmentalised one in which the marketing people specify customer preferences; design engineers devise specifications; process engineers translate the specification into manufacturing plans; plant engineers configure machinery and work practices to the manufacturing plans; and once production starts, line workers execute the plant engineers’ instructions. By

<sup>33</sup> Stanley Fish, ‘Still Wrong After All These Years’ in *Doing What Comes Naturally* (Baltimore, MD: Johns Hopkins University Press, 1990) 356–71. As Holmes put it, ‘historic continuity with the past is not a duty, it is only a necessity’. Oliver Wendell Holmes, ‘Learning and Science’ in *Collected Papers* (New York: Harcourt Brace, 1920) 13839.

<sup>34</sup>

For example, the plant has had a persistent problem with a bracket on the brake pedal sub-assembly to which the cables for both the cruise control and power brakes are attached. The bracket often moves when the cruise control is used, resulting in misadjustment of the breaks. [T]he problem was a poor design—that the cruise control and brake cables shouldn’t be attached to the same bracket, and that the bracket was in a bad location. [But to add or reposition a bracket would require] a long struggle with Detroit. (MacDuffie, n 18 above, at 484).

contrast, the Toyota style is concurrent. Teams with members drawn from all these fields work together to consider simultaneously customer needs, cost, technical performance and manufacturability. As the time for manufacture approaches, 'pilot teams' are recruited from the shop floor to work on the specification of jobs for manufacture of the new models, and they sometimes make suggestions that result in design revision. Pilot versions of the new models are sometimes built on assembly lines that are concurrently building the old models in order to facilitate a quick transition.<sup>35</sup>

In all these ways, TPS blurs the distinction between the prescriptive and the instrumental, or the elaboration of goals and the specification of means.

*E. Individualistic and retrospective concerns are sublimated or bracketed.* TPS is intensely preoccupied with mistakes, but its orientation toward them is prospective and collective. The issue is how the group will fix the problem. In his study of three auto plants, MacDuffie found the two American-owned ones preoccupied with determining who should be held responsible for mistakes. By contrast, at the Honda plant which had more fully embraced TPS, '[t]he accounting system is deliberately designed to minimise the time spent figuring out who's to blame'.<sup>36</sup>

In the TPS perspective, problems are more likely to result from defects in system design than from blameworthy deviance on the part of particular workers. Assigning responsibility consumes unproductive time and effort and encourages recrimination that undermines solidarity and willingness to volunteer information about mistakes. TPS thus presumes that 'a problem with our product is a problem for the whole company, not an individual or department'.<sup>37</sup>

The difference between 'root cause analysis' and typical common law liability analysis is revealing. The latter stops when it finds blameworthiness, and it usually measures blameworthiness by departure from established norms. On the other hand, root cause analysis is primarily concerned with how the norms could be improved. Moreover, when root cause analysis discovers violations of established norms, it does not stop. It goes on to consider whether the violations are symptomatic of some background condition that might be improved. Operator negligence, for example, might be a symptom of deficient training, or workplace distractions, or excessively long shifts. (A common law court might go on to consider these systemic factors if the plaintiff could make a case that they departed from established norms, but whether it did so would depend on the plaintiff's individual concerns, such as whether institutional liability would be necessary in order for a judgment to be collectable, rather than the potential benefits to the system.)

<sup>35</sup> Womack, Jones, and Roos, n 16 above, at 138–68; Adler, 'Flexibility vs. Efficiency' n 17 above.

<sup>36</sup> *Ibid* at 493.

<sup>37</sup> *Ibid*.

54 *William H Simon*

The TPS system emphasises intangible incentives, such as peer respect, over material ones, and group incentives over individual ones. Contingent compensation is often on a company, plant, or group basis, and rarely on an individual one.

Nevertheless, we've seen that TPS is nested in larger structures that contain norms that more approximate the themes of mainstream jurisprudence. I noted above that there are background clusters of secondary norms that allocate dispute resolution authority. There are also background clusters that are responsive to concerns about individual fairness and incentives. Three are especially notable.<sup>38</sup>

First, in Japanese firms, TPS is typically accompanied by a parallel personnel system that operates in critical respects more hierarchically and informally than the production system. This system provides for the ranking of employees, mostly by seniority but also on the basis of informal supervisory assessments of performance. This ranking has a small influence on compensation and a larger one on promotions. Supervisory judgment is checked both by administrative review and by the possibility of challenge through a union grievance process. In some companies, employees dissatisfied with their supervisors can easily transfer to a different unit.

Second, adoption of TPS typically correlates with a relatively high degree of employment security, and TPS proponents often insist that the latter is a condition of the former. Workers are likely to be more cooperative when productivity increases do not threaten their jobs and when their pay and security do not depend on their particular position in the company. Job security encourages long-term identification with the company, which in turn encourages voluntary effort.

Third, worker and employer in the Toyota-style firm typically bind themselves to each other through reciprocal investments. The firm invests in the worker by providing the more extensive training in general-purpose skills that TPS requires relative to traditional manufacturing. The worker invests in the firm by accepting back-loaded, seniority-based compensation. Each has an incentive to satisfy the other and preserve the relation.

Thus, individual fairness and incentive issues have not been ignored in TPS. However, their salience has been reduced through bracketing, muting and deferral. Fairness issues are bracketed simply by removing them from the shop floor to a separate department with separate personnel. They are muted by the bonding practices that mean that the stakes in any individual claim arising from conduct in the production process will often be dwarfed by the shared interest in preserving the relationship. And they are deferred by a system that provides rewards for success. To be sure, fairness issues can potentially arise over the division of the fruits of success. But it is a common experience that collaborators find it easier to allocate the benefits of

<sup>38</sup> See generally Masahiko Aoki, *Information, Incentives, and Bargaining in the Japanese Economy* (Cambridge: Cambridge University Press, 1988) 49–98.

success than the burdens of failure. By promising contingent group compensation, the system may reduce individual fairness claims by deferring some of the potential for them to a future point where, if success has occurred, the pressure to press them will have relaxed.

The bracketing, muting and deferral of individual incentive and fairness issues make Toyota Jurisprudence better able than its mainstream counterpart to address organisational actors. The basic units of Toyota Jurisprudence are groups rather than individuals. The groups, however, have fluid identities that are reconstituted continuously in the problem-solving process. Problem-solving groups may cross departmental divisions and even legal boundaries between firms. The participants are motivated by collective incentives, both material and immaterial ones, such as solidarity and craft pride. A system of rewards and sanctions calibrated to individual performance would generate centrifugal pressures that would fragment groups. Even a retrospective system preoccupied with past group performance would be disruptive because it could not take account of the fluid re-composition of groups; the groups currently being rewarded or sanctioned might not have the same composition as they did at the time of the relevant past conduct. On the other hand, the focus on collective goals and incentives makes it possible for Toyota Jurisprudence to treat groups as central.

### 3. TOYOTA JURISPRUDENCE AND PUBLIC LAW INNOVATION

The Toyota perspective resonates with some important recent trends in American public law that my Columbia colleagues call 'experimentalist' and that Carey Coglianese and David Lazer call 'management-based regulation'.<sup>39</sup> The trends are evident in several areas of health and safety regulation. They are also prominent in the approach to school reform partially codified in the No Child Left Behind Act.

Health-and-safety programmes that exemplify these tendencies include the Hazard Analysis and Critical Control Point (HAACP) food safety programme of the Department of Agriculture, the Process Safety Management Program for hazardous substances of the Occupational Safety and Health Administration (OSHA), the safety regime for nuclear power plants overseen by the Nuclear Regulatory Commission and the Institute of Nuclear Plant Operations (INPO), and the Massachusetts Toxics Use Reduction Act (TURA).<sup>40</sup>

<sup>39</sup> Dorf and Sabel, n 2 above; Cary Coglianese and David Lazer, 'Management-Based Regulation: Prescribing Private Management to Achieve Public Goals' (2003) 37 *Law and Society Review* 601.

<sup>40</sup> See Coglianese and Lazer, previous n, at 696–700 (HAACP and Process Safety); Bradley C Karkkainen, 'Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?' (2001) 89 *Georgetown Law Journal* 257, 354–56 (TURA); Joseph Rees, *Hostages to Each Other: The Transformation of Nuclear Safety Since Three Mile Island* (Princeton, NJ: Princeton University Press, 2003) (NRC and INPO).

56 *William H Simon*

These regimes proceed by requiring regulated actors to identify hazards of particular kinds and to formulate their own plans for dealing with them. The plans must be based on research (not necessarily the actor's own) showing the efficacy of the measures prescribed. The plans must be periodically reviewed and revised in the light of experience. They must specify tolerances and indicators to define satisfactory performance and corrective measures when the performance is unsatisfactory. Actors are audited for the adequacy of their plans and for compliance with them. The regimes (or related programmes) specify mandatory practice or performance standards, but the actor will also be held to the standards its own plans provide, which are expected to be higher than the mandatory ones in many respects.

Some firms will be motivated to exceed minimum standards because such performance coincides with other firm goals, such as minimising cost or smoothing product flow, or because they are subject to distinctive pressures for better performance from workers, customers or neighbouring residents. As these firms improve measurably, they demonstrate possibilities for similar firms, and give the stakeholders in these other firms and the regulators a basis for demanding more of them.

Some compliance judgments may be binary (turning on whether the actor has met a fixed threshold), but others will rank the actor on a scale that permits comparison both with its own past performance with the performances of peers. Minimum standards may ratchet up as leading performers raise expectations. Or standards may require a minimum quantum of improvement over past scores. The regulator or an industry trade association collects and disseminates information about 'best practices' associated with the highest performances. The remedial aspects of the programme tend to be less punitive than those associated with command-and-control regulation. Lagging actors are asked to diagnose themselves and formulate improvement plans, perhaps with technical assistance from the regulator. They are subjected to increased monitoring. Publication of their performance ratings may generate shaming pressures within the industry or more tangible responses from customers or investors. Continued severe performance failure may result in more punitive intervention, but the typical pattern of intervention is less one of gradually increasing harshness, and more one of prolonged cooperative intervention followed, in the most intractable cases, by complete exclusion of the actor or forced restructuring.

'New accountability' school governance programmes as pioneered in Texas, Kentucky and North Carolina, and partly mandated by the federal No Child Left Behind Act have a similar structure.<sup>41</sup> The federal statute does not prescribe substantive standards or practices. Instead, it requires states to set their own goals for their schools and then to develop measures

<sup>41</sup> See James Liebman and Charles Sabel, 'A Public Laboratory Dewey Barely Imagined: The Emerging Model of School Governance and Legal Reform' (2003) 28 *New York University Journal of Law and Social Change* 183.

to make schools accountable for efforts to fulfil the goals. State systems must develop standardised tests to measure attainment of their own goals, and must also test their students on a uniform national test. They must report the performance of each school overall and with respect to ethnic minorities, economically disadvantaged, and disabled sub-groups. All schools are expected to annually review their performances and plan for improvement.

Local education agencies must publish annual 'report cards' comparing the performance of each school with schools across the state with socio-economically comparable student populations. The agencies must provide reward to high performers and technical assistance to laggards in drafting and implementing improvement plans. Where low performing schools fail to improve, students acquire rights to transfer to other schools or to use their share of federal support for the school to purchase educational services of their choice. Persistent poor performance ultimately requires that the schools be reconstituted with new management.

The main contours of these programmes strongly resemble key features of TPS.

#### *A. Purposes*

Important features of these regimes seem designed neither to resolve disputes nor to vindicate accepted values, but to induce learning and innovation. These include the features that oblige the actors to define and justify their own standards and practices, to make them public in a way that permits others to observe their experience, and to continuously revise them in the light of shared experience. The regimes are designed to induce performance by some actors at a higher level than anything the norms currently specify.

In addition to continuous revision, the rules require special responses to problems—indicators of failure or potential for improvement in the system. There are three types of such indicators: First, relative performance scores on measures used in periodic audits, for example, the number of unplanned shutdowns at a power plant or student scores on a standardised test. Second, observations or test results indicating failure to comply with applicable practice norms, including both agency mandatory rules and the firm's plans. And third, abnormal events, near misses and breakdowns—for example, the accidental discharge of a hazardous substance in the workplace.

Poor performance on any of these indicators may lead to sanctions and dispute-resolution procedures. But the most prominent and characteristic response the systems prescribe involves neither. The firm, often in collaboration with the regulator, must investigate and develop changes designed to improve its performance. The demand is for improvement, not necessarily for vindication of a pre-existing norm. Minimum standards have to be met, but the demand for improvement may require considerably more. The ethos of these regimes is hostile to the 'compliance orientation' that takes

58 *William H Simon*

meeting fixed standards as adequate. Moreover, the decision-making procedure contemplated for investigation and reform is consensus both within and between the regulatory agency and the firm.

Of course, these systems are nested in larger structures that include rules with mandatory standards requiring performance in accordance with current norms and procedures for non-consensual dispute resolution. A system in which these mandatory standards and non-consensual procedures played a dominant role, however, would be regarded by its designers as a failure. Nevertheless, a theorist might want to reserve the terms 'law' or 'legal system' for these more traditional norms and insist that the Toyota-style norms represented a non-legal form of social order. There would be significant costs to doing so, however. At an abstract level, the Toyota norms have significant kinship with the more traditional ones. They are both public systems of deliberate social order based on rules and related norms. Moreover, the Toyota-style systems are typically regarded by their designers and participants as alternatives to the more traditional systems.

*B. Rules and standards*

In all these systems, we find an emphasis on specified norms coupled with a duty to revise them continuously. Plans and performance measures should be as specific as possible. Hazard Analysis and Critical Control Points food safety plans, for example, must specify safety practices in detail, precise tolerances for problem indicators such as salmonella, and precise corrective actions when tolerances are exceeded. Yet, corrective action is a matter of re-assessment as much as enforcement. The norms should be re-written when experience indicates they can be improved.<sup>42</sup>

As in TPS, the systems sometimes give rank-and-file workers the duty to provoke reconsideration of a practice norm whenever its application would defeat its purposes. Institute of Nuclear Power Operation standards provide:

If the individual actually performing the activity cannot or believes he should not follow the procedure governing that activity as written, he shall place the system/component into a stable and safe condition and inform the responsible supervisor. Situations such as this could occur if the procedure is found to be inadequate for the intended task, if unexpected results occur, or if two more procedures governing the activity conflict. The supervisor shall resolve the discrepancy in the procedure by either [determining that the procedure is in fact adequate

<sup>42</sup> US Dept. of Agriculture, Food Safety and Inspection Service, 'Pathogen Reduction; Hazard Analysis and Critical Control Point (HAACP) Systems' 61 Fed Reg 38805 38816 (25 July 1996) ('[I]n the course of corrective action, t]he HAACP plan itself might require modification, perhaps in the form of a new critical limit, or of an additional [Critical Control Point].')

or] submitting a procedure change... (no further procedural steps shall be accomplished until the procedure change is approved).<sup>43</sup>

In new accountability school reform, performance is measured precisely on standardised tests, but the tests are continuously revised as understanding of goals changes or as knowledge of how to measure goals improves. Administrators and teachers make and revise detailed plans for improvement at the district, school, and classroom levels.<sup>44</sup>

### C. Interdisciplinary teams

The regulatory regimes contemplate problem-solving by interdisciplinary teams. For example, OSHA recommends for the required investigation of an unplanned release of hazardous chemicals:

A multi-disciplinary team is better able to gather the facts of the event and to analyze them and develop plausible scenarios as to what happened, and why. Team members should be selected on the basis of their training, knowledge and ability to contribute to a team effort to fully investigate the incident. Employees in the process area where the incident occurred should be consulted, interviewed or made a member of the team. Their knowledge of the events form a significant set of facts about the incident which occurred.<sup>45</sup>

The Texas education statute requires site-based management teams with representatives of administrators, teachers, campus staff, parents and community members. Reformers emphasise collaboration among teachers both within and across departments and between teachers and administrators. The new system generates two distinctive pressures for collaboration. First,

<sup>43</sup> Institute of Nuclear Power Plant Operation, *Good Practice, Conduct of Operations* (July 1984) 18–19 (quoted in Rees, n 40 above, at 82). The rule continues:

In cases of emergency when procedures are inadequate for the situation, plant operations personnel are directed to take such action as necessary to minimize personnel injury and damage to the plant, to return the plant to a stable, safe condition, to protect the health and safety of the general public and the personnel on site. These actions shall be documented and, if appropriate, incorporated into a revision of the affected procedure.

<sup>44</sup> Liebman and Sabel, n 41 above, at 249–50.

<sup>45</sup> 29 CFR 1910.119, Appendix C, para 12. An earlier version of the rule specifically recommended that each team include:

- A third-line or higher supervisor from the section where the incident occurred;
- Personnel from an area not involved in the incident;
- An engineering and/or maintenance supervisor;
- The safety supervisor;
- A first-line supervisor from the affected area;
- Occupational health/environmental personnel;
- Appropriate wage personnel (i.e., operators, mechanics, technicians); and
- Research and/or technical personnel.

(US Dept of Labor, Occupational Health and Safety Administration, Proposed Rule, 55 Fed Reg 29, 150, 29172)



## 60 *William H Simon*

performance assessment with standardised tests requires each school to teach a standard set of skills and knowledge, and this means that within the school, teachers must coordinate their teaching to ensure the proper coverage. Second, the aspiration to respond quickly to information gleaned in the course of the school year requires various forms of collaboration. Students or teachers may need to be re-assigned; new teacher training may need to be arranged; curriculum may need to be adjusted to increase emphasis on some skills and reduce emphasis on others.<sup>46</sup>

### *D. Destabilisation*

These systems incorporate benchmarking and root cause analysis practices.

The Institute of Nuclear Power Operation disseminates rankings of nuclear plants based on performance indicators. Plants are distributed across five categories ranging from 'Excellent' to 'Marginal'.<sup>47</sup> The Texas Educational Accountability System publishes a similar set of rankings of Texas schools grouped in socio-economically homogeneous peer groups.<sup>48</sup> The practices of the most successful performers are often publicised. Observers of both regimes find that these activities have a powerful influence both in motivating performance through honour and shame and in diffusing knowledge about 'best practices'.

Corrective action plans are based on root cause analysis. They are most elaborate in nuclear safety, where 'significant operating event' investigations often lead to changes in operations, maintenance, training and administration.<sup>49</sup>

A recent controversy involved in the implementation of the HAACP system by the Food Safety and Inspection Service parallels issues MacDuffie identified in auto plant TPS implementation and illustrates the stakes in root cause. Meat tainted by e-coli H 057 was found at Gallison Wholesale Meat Company in Montana in 2002. The Food Safety and Inspection Service inspectors encouraged a recall and demanded corrective action from Gallison. However, the inspectors failed to make efforts to trace the meat back to Gallison's suppliers to see if the taint had originated with them. Operating on traditional assumptions, the inspectors assumed that it was

<sup>46</sup> See Charles A Dana Center, *Driven to Succeed: High Performing, High Poverty, Turnaround Middle Schools* (Austin, TX: Charles A Dana Center, 2002), available at <[http://www.utdanacenter.org/downloads/products/driven/ms\\_vol1.pdf](http://www.utdanacenter.org/downloads/products/driven/ms_vol1.pdf)>

Judges are not central participants in the regimes considered here. Thus, the regimes do not alter the judicial role. An example of a reform in the spirit of TPS in which courts are central and judges become members of interdisciplinary teams—highly controversial in part for this reason—is the drug court. See Dorf and Sabel, n 11 above.

<sup>47</sup> Rees, n 40 above, at 98–10.

<sup>48</sup> Liebman and Sabel, n 41 above, at 241.

<sup>49</sup> See Rees, n 40 above, at 126–50.



the wholesaler's responsibility to guarantee the safety of the meat, whether the taint originated from its own processes or those of its suppliers. After a critical report from the Department of Agriculture's Inspector General insisted that HAACP required tracing back, FSIS revised its regulations to require it in such circumstances. Now problems such as this potentially implicate the entire supply chain.<sup>50</sup>

#### *E. Prospectivity and remediation*

These regimes are oriented toward problem solving, as opposed to blaming or punishment. The search for causes is more a search for improvement opportunities than an effort to assign blame. One finds a combination of soft incentives—shaming and honour—with extremely hard ones for persistently poor performers. When improvement seems unlikely, the regimes prescribe management change, facility shutdown or license revocation.

But purely punitive sanctions are typically not a direct part of the regime. They stand in the background, threatening intervention in cases of deliberate malfeasance, especially withholding and falsification of data. The system designers tend to bracket or sublimate such concerns. Punitive interventions are relegated to separate sub-systems of norms, not typically referred to in the regimes' self-descriptions, and often assigned to different personnel. This de-emphasis reflects assumptions that non-compliance is more often the result of incapacity than malfeasance and that the prospect of punishment deters the cooperation on which the regimes depend.<sup>51</sup>

#### *F. Death penalty administration*

Regulatory approaches that resonate with TPS may not be promising for all spheres, but their range has yet to be mapped. There's a tendency to think they are not well suited to issues that are highly charged morally. While there is something to this contention, consideration of Jim Liebman's proposals for reform of death penalty administration suggests that it is overstated.<sup>52</sup>

Liebman's proposals emerged from his extensive empirical study showing that, in about two-thirds of all cases imposing the death penalty, reviewing courts find errors sufficient to warrant post-conviction relief. This relief

<sup>50</sup> US Dept. of Agriculture, Office of Inspector General, Great Plains Region, 'Audit Report: Food Safety and Inspection Service Oversight of Production Process and Recall at Conagra Plan (Establishment 969)' (Sept. 2003) at 43–48.

<sup>51</sup> See, eg, 29 CFR 1910 119, App C, para 12:

The cooperation of employees is essential to an effective [workplace hazardous substance] incident investigation. The focus of the investigation should be to obtain facts, and not to place blame.

<sup>52</sup> James Liebman, 'The Overproduction of Death' (2000) 100 *Columbia Law Review* 2030.

62 *William H Simon*

comes in a procedure that typically takes many years, often decades, and includes multiple rounds of review in each of the state and federal court systems.

Liebman showed how the current system is structured to 'overproduce' death sentences: Prosecutors get political capital (a reputation for toughness on crime) and procedural advantages (the opportunity to death-qualify the jury) from seeking the death penalty. Elected state court judges face a political price for checking their excesses. Neither prosecutors nor judges face much reputational cost when wrongful convictions for which they are responsible are identified because the system makes no effort to connect the mistakes with those who are responsible for them and because mistakes are discovered so long after they are made that the responsible actors are likely to have moved on. Prosecutors, judges, and juries take their responsibilities toward the accused casually because they believe (with some plausibility) that any errors they make against the accused will be remedied post-conviction. Publicly provided defence resources at trial are inadequate. Although there is a talented private bar specialising in death penalty cases, it is forced by its small size and limited resources to focus on post-conviction challenges, and hence does nothing to check the tendency of the system to overproduction at the charging and trial stages.

Liebman's reform proposal involves a curtailment of the present elaborate post-conviction review procedures in return for strengthening protections for the accused at the stages of the trial and the initial state appellate review. It includes many specific reforms long advocated by defenders, including enhanced funding and heightened qualifications for defence lawyers, open access by the defence to the prosecution's files, a requirement that confession evidence be videotaped and a prohibition on death-qualifying juries in the guilt phase. Two aspects of the package, however, are more distinctive.

First, prosecutors must explicitly justify their charging decisions, and state judges must review sentences for consistency with both the prosecutor's rationales and decisions in other cases. The prosecutorial justification comes in a statement filed at least 120 days before announcing a decision to seek the death penalty. Comparative proportionality review is to be undertaken post-conviction by the state appellate courts.

Second, data with respect to mistakes by courts, prosecutors and defence attorneys must be compiled in a form that permits comparison across personnel and jurisdictions and be disseminated.

These proposals resonate with the Toyota perspective.

First, their goals are partly to induce learning. To be sure, the criminal justice system as a whole is necessarily strongly focused on dispute resolution and the vindication of accepted values. As long as the defendant denies allegations of the indictment, dispute resolution is critical. And just punishment is punishment that conforms to accepted values.

Nevertheless, it appears that, even among people who accept the justice of the death penalty in the abstract, there is no consensus as to when it should be imposed. The pattern of charging and sentencing in capital cases

strikes most observers as erratic and as an affront to widely accepted values of equal treatment. Liebman's proposed requirements with respect to the explanation of charges by prosecutors and proportionality review by first-stage appellate courts are designed to induce reflection and to generate information that makes more consistent and thoughtful judgments possible. The proposals are not an effort to vindicate a consensus, but to make it possible for one to emerge through the public deliberative efforts of different courts, each taking account of the others decisions and reasoning.

Second, the proposals are consistent with a Toyota-style response to the dialectic of rules and standards that has played out in an especially troubling way in this area. Sentencing decisions under discretionary standards tend to be erratic or discriminatory. When reformers respond with rule-like sentencing grids in the style of the Federal Sentencing Guidelines, they often compel arbitrary results in cases they fail to anticipate or provide for. (For years, the Supreme Court could not make up its mind whether due process in death sentencing was more offended by the looseness of discretion or the rigidity of rules.) The Toyota approach suggests the possibility of continuously revisable grids. The grids would have only presumptive force. Departures would be permitted if the decision maker could give reasons for them, and each departure would be an occasion for consideration (perhaps by a sentencing commission) of whether the grids could be improved.

Third, although the element of collective decision making is not salient in the reforms, there is a tendency of the requirements of justification to encourage more engagement within prosecutorial offices and between trial judges. The need to publicly justify charging decision and the appellate review of them for consistency will encourage prosecutorial staffs to develop their practices in a more open and coordinated fashion. While trial judges will not directly deliberate with each other, they will have more reason and opportunity to take account of each other's decisions in sentencing in order to achieve consistency.

Fourth, Liebman's proposal to generate and publish comparative error data with respect to prosecutors and courts is an attempt to generate the destabilising pressures of benchmarking and root cause analysis. The hope is that rankings will encourage emulation of the most successful performers and stigmatise the laggards in ways that create pressures to reform. The attribution of the errors requires a kind of root cause analysis. The overall tendency of the proposals, which limit post-conviction review in return for more safeguards in the earlier process, is to reduce reliance on ad hoc end-of-process adjustment ('rework') in order induce improvements in early-stage routine performance.

Fifth, by shifting attention from procedures for the case-by-case rectification of past errors to systemic reforms for reducing errors, the proposals have a less retrospective and individualist orientation than conventional discourse. Capital punishment will always be fundamentally a matter of corrective justice. But doing corrective justice requires institutional capacities that cannot be developed only case by case.

## 4. A JURISPRUDENCE OF PROBLEM-SOLVING

The jurisprudence suggested by TPS is a jurisprudence of problem-solving. Such a jurisprudence might overlap substantially with rights-and-principles and law-and-economics legal theory. It could make ample use of the structures of moral discourse elaborated by the rights-and-principles theorists and of both the empirical and analytical techniques of the economists. It would be considerably less troubled than either of these groups tends to be about the tensions of its precepts with those of the other, since it wouldn't take either as a contender for a comprehensive theory.

However, rights-and-principles theory is primarily preoccupied with *interpretation*. A jurisprudence of problem-solving, by contrast, would be focused more on *deliberation*. Law-and-economics is primarily a theory of *optimisation* of known preferences in known circumstances. A jurisprudence of problem-solving, by contrast, would be focused more on *discovery*—on the ways in which people can deepen their understanding of both their goals and the possibilities of realising them.

John Dewey's style of pragmatism offers a philosophical basis for theories of problem-solving, but it tends to be disappointingly vague about practical application.<sup>53</sup> The engineering ideas and practices associated with TPS and reflected in recent public law developments might usefully complement the pragmatist approach and contribute to a problem-solving jurisprudence.

The potential benefits of a jurisprudence of problem-solving extend beyond the capacity to come to terms with recent developments and regulation and public service delivery. For, as I suggested above, a substantial range of longstanding concerns of the American legal system do not fit well with the preoccupations of mainstream legal theory. One set of these concerns is substantive; it is typified by unintentional torts and traditional health and safety regulation—laws that regulate the unintended consequences of generally beneficial conduct. A major fraction of the conduct prohibited by these regimes is not plausibly moralised and cannot be described or identified precisely for deterrence purposes. The other set of concerns involves organisational liability. Mainstream legal theory has been developed largely with reference to individuals; it has very few resources for dealing with the distinctive aspects of organisational liability. Toyota jurisprudence, which brackets moralism, acknowledges uncertainty, and takes collective action as paradigmatic, has potentially important contributions to make in both areas.

<sup>53</sup> See Dewey, n 2 above.