9

Minimally sufficient punishment

Key propositions

- For unusually oppressed people, like First Nations Australians, punishment is not a minor facet of their domination, but is central to the dynamics of domination.
- Deterrence works best when it is progressively reduced with the aim of growing freedom as nondomination. It must be reduced to the lowest level of deterrence that can avert the escalation of crime.
- Which strategy works best at crime control is not the most important question for criminology. More important is which meta-strategy is best in a given situation; which strategy for sequencing strategies best reduces crime and domination? Deterrence contributes best to meta-strategy design when punishment is low and decreasing, but detection is perceived to have high certainty, and escalation is seen as inexorable without desistance.
- Deterrence works best when escalation of deterrence is combined with escalation of social support to help offenders take paths away from punishment.
- Successful crime prevention persuades offenders that trouble hangs inexorably over their head, but caring people will support them to avert it.
- Freedom depends on escalating social support until desistance from domination is consolidated.

- Deterrence above minimal sufficiency blunts deterrence. Anything more than a minimally sufficient frequency of escalation to deterrence blunts deterrence for future cases.
- A preference for restorative justice over deterrence sharpens future deterrence of crimes of the powerless and, more surprisingly, future deterrence of crimes of the powerful.
- Deterrence usually fails because the criminal justice system always faces a system capacity crisis that is at its worst when and where the crime rate is worst. Responsive escalation helps solve the system capacity crisis by motivating most punishment to be self-punishment and most prevention to be self-prevention. Responsive regulation rations punishment to cases where ethical appeals for remorse, apology, reparation and self-prevention of future offending do not work.
- When intrinsic motivation to comply with the law is kept intact, responsive regulatory enforcement chooses not to crowd out intrinsic motivation with extrinsic threats.
- Responsive enforcement has a dynamic design to ensure that game-playing to avoid legal obligations inexorably produces escalation to deterrence and then incapacitation.
- Deterrence works best when it focuses on a line that should never be crossed after an announcement date, followed by progressive lifting of that line, raising our expectations of responsible corporate and individual citizens.
- Law enforcement works best when it averts stigmatisation, while communicating the shamefulness of predatory crime.
- Freedom is maximised when the structural punitiveness of the system is gradually reduced until punishment gets so low that insufficiency of punishment increases crime.
- Minimally sufficient punishment allows the least punitive societies to close most of their prisons while meeting the UN Standard Minimum Rules for the Treatment of Prisoners (the 'Nelson Mandela rules').
- This is best done with pride and publicity that educate citizens about why 'jailing is failing'.¹

1 'Jailing is failing' is the campaign message of the Justice Reform Initiative in Australia, of which I am proud to be an ACT patron.

Dynamics of just enough deterrence

Threats to freedom must be deterred, but deterrence is overrated compared with other crime-control tools discussed in this book. The next chapter considers how incapacitation, especially self-incapacitation, is a more useful doctrine of prevention than deterrence. It also argues that when incapacitation is not mainly in the form of imprisonment, it can be a less dominating doctrine than deterrence. Captivity is not the best circumstance for cultivating capabilities for freedom. Most Aboriginal and Torres Strait Islander people in Australia are arrested by the police during their youth, often with deeply stigmatising consequences during their schooling and as they attempt to get their first job. This is something that happens to quite a small minority of white Australians. Aboriginal and Torres Strait Islander people are even more overrepresented in prison populations than African Americans in the United States. The intersection of a criminal record and race makes it impossible for a large proportion of the Indigenous population to sustain employment. Overreliance on deterrence and prison is thus central to the domination suffered by disadvantaged minorities. My contention is that there is little hope of tackling racial inequality without emptying prisons of 90 per cent or more of their occupants.

This chapter is about how to achieve just enough dynamic deterrence to secure freedom through minimally sufficient deterrence. For more, see Braithwaite (2008, 2018), on which this chapter expands and draws heavily.

Dangers exist in maximalist approaches to deterrence and in minimalist ones (such as that restorative justice can replace punishment). A minimal sufficiency strategy aims to avert these dangers. The objectives are to convince people that the webs of relationships within which they live mean that lawbreaking will ultimately lead to bad outcomes. These webs of relationships can also persuade offenders that predatory crime is simply wrong. The pitch to offenders is to abandon criminal careers because doing so assures desistance of the state and of loved ones from increasing intrusions in their lives. Social support lays a path to desistance that is also a path to freedom for the offender. Alternative support and control strategies should be attempted until desistance finally occurs. Communities can be helped to understand that this is how minimally sufficient deterrence works. By relying on layered strategies, this approach takes deterrence

theory on to the terrain of complexity theory. It integrates approaches based on social support, recovery capital and collective efficacy, dynamic concentration of deterrence, restorative justice, responsive regulation, responsivity and indirect reciprocity. Deterrence is desolate in its failure as a criminological doctrine because of its rejection of complexity in favour of simple theories such as rational choice.

Some criminologists are inclined to ask why any role would be given to deterrence. The evidence for the power of deterrence in reducing crime is thin, after all (Nagin 2013; Chalfin and McCrary 2017; Tonry 2018). There are three answers to why deterrence should retain a significant role. One is that a good meta-strategy for crime control achieves strength through the convergence of weaknesses: deterrence can help to motivate crime-control strategies that are more effective. Second, deterrence is one weak strategy that can be tried after various less weak strategies have failed, strengthening the efficacy of a complex mix that is tied together as a bricolage of strategies. In other words, deterrence does work, but rarely on its own and mainly when deterrence is woven into a regulatory mix. Third, when deterrence of a specific offender fails, it might slightly strengthen the general deterrence of other offenders.

Sometimes a tax audit teaches a corporate chief financial officer more about what they can get away with in an audit than what they cannot. Sometimes punishment or the threat of punishment provokes defiant reactions that can make crime more, not less, likely. For most values of relevant variables, defiance effects exceed deterrence, but there are some contexts in which specific deterrence exceeds defiance. For these reasons, deterrence minimalism is rejected in favour of minimally sufficient deterrence.

Deterrence maximalism is also rejected. Zero tolerance and other political slogans that go to deterrence maximalism are common. They are doubtless helpful in some kinds of election campaigns but are rarely taken seriously by scholars who understand the evidence. Deterrence can never be the main game of crime control. Even so, it is reckless to fail to develop a view of the constructive role deterrence must have in crime prevention. The data on the limited effectiveness of deterrence and the cost of prisons (Nagin 2013; Durlauf and Nagin 2011a, 2011b; Travis et al. 2014; Petrich et al. 2021) demand disinvestment from locking up offenders. This is a cornerstone of 'justice reinvestment': disinvestment from prison

and reinvestment in evidence-based social support pathways. It is easy to dismiss the prescriptions of maximalists who push for sentences that are as long as the political process can drive them.

Likewise, it is easy to dismiss maximising the shame aimed at offenders. While there are criminologists who argue that shame has power in crime control (Braithwaite 1989), these scholars do not advocate maximising the denunciation directed at offenders. The evidence is that this strategy leads to stigmatisation, which makes crime worse (Ahmed et al. 2001: 3–72; Braithwaite 2020c). Moreover, we now know that healthy pride management may be quantitatively as important as, or more important than, healthy shame management (Ahmed and Braithwaite 2006; Maruna 2001; Best et al. 2016). Intentionally directing unhealthy shame at offenders may crowd out healthy pride from the encounter. There is unhealthy shame that increases crime and healthy shame acknowledgement that helps prevent crime and repair harms. Likewise, there is unhealthy pride that fosters crime by vaunting superiority over others, and there is humble pride in doing things well with others that is vital to crime prevention—often via pride in the identity of being a law-abiding citizen who cares about the suffering of others (Ahmed and Braithwaite 2006). Humble pride in citizenship obligations is a building block of freedom. What is needed is an approach to deterrence that does not crowd out healthy shame acknowledgement and healthy pride in a law-abiding self. Indeed, a strategy that nurtures them is required. A virtue of minimally sufficient deterrence is that it minimises that stigmatic crowding out that is inherent in deterrence that brutalises.

Progressives who seek to minimise the quantum of fear or shame that criminal processes invoke are also a danger. For example, this chapter contends that it is dangerous to regard restorative justice as an abolitionist prescription that eliminates the need for punishment and deterrence. Restorative justice is a strategy to give an opportunity to all the stakeholders in a crime to participate in a process that discusses who has been harmed, who has needs and what might be done to repair those harms and meet those needs (Zehr 2015). It is about the idea that because crime hurts, justice should heal. A naive aspect of the view that restorative justice eliminates the need for deterrence is denial of the reality that if we gave criminal offenders the choice between agreeing to meet their victim in a restorative circle to discuss repair of the harm or doing nothing and forgetting about it, most offenders would opt to forget it. Offenders mostly agree to participate in restorative justice because the alternative

has deterrent elements. We see from this that a useful role for deterrence is to motivate engagement with something that is more effective than deterrence: restorative justice and the rehabilitative and preventative measures for which restorative justice is a delivery vehicle. Indeed, wise integration of restorative justice and deterrence allows restorative justice to strengthen the preventive power of deterrence, in addition to allowing deterrence to strengthen restorative justice. This illustrates the ambition of identifying a good meta-strategy for crime control that achieves strength from the convergence of weaknesses: in this example, strength from the convergence of the weaknesses of deterrence and restorative justice.

Some restorative justice advocates are reluctant to see a positive role for shame or are minimalists about shame (for example, Maxwell and Morris 2002). A society in which rape, violence and corporate crime are minimally shameful will be a society with high rates of rape, violence and corporate crime (Braithwaite 1989, 1995, 2020c). Hence, it is also imperative to diagnose what minimal sufficiency of the right kind of shame might mean.

There are alternative paths between maximalist and minimalist approaches: a minimal sufficiency strategy of deterrence guides us towards them. Minimum deterrence and minimum shame are inferior to minimal sufficiency, which means just enough of the right kinds of deterrence and the right kinds of shame (which shun stigma as a criminal law doctrine or an objective of sentencing). Deterrence and shaming are more effective when combined with a dynamic theory of social supports partly because supports render shame more reintegrative. Communicating the shamefulness of predatory crime is more effective when combined with the reintegration of offenders (Braithwaite 1989). As with deterring crime, the evidence is that deterring warfare works better when armed fighters are simultaneously shown the costs of killing and shown a supportive peace dividend that benefits them, their family and their community (Braithwaite and D'Costa 2018: Ch. 3; Toft 2010). This is why there is strong evidence that peacekeepers prevent war when they do multidimensional peacebuilding that supportively delivers peace dividends (Doyle and Sambanis 2006; Walter et al. 2020). It is also why simplistic strategies of deterrence maximalism by backing threats to foes with investment in ever more battalions and bombs get countries into more, not less, war.

In business regulation, we have seen that perhaps the most cost-effective strategy is informal praise by inspectors when companies improve because a statistically significant improvement in compliance is achieved at near zero cost (Chapter 5; Makkai and Braithwaite 1993), along with benefits to morale, motivational postures (Braithwaite 2009a) and regulatory legitimacy. Data located at a radically different level from street-level encounters of inspectors also support the pivotal role of praise compared with shame.

The meta-analyses show consistent, statistically significant positive effects of the reporting of environmental stewardship on the financial performance of large firms, and even bigger effects for small firms (Dixon-Fowler et al. 2013). The effect sizes are not huge, but they are sufficient to make it generally rational for firms of all sizes to work hard at being good environmental citizens. The empirical work of Amato and Amato (2012) suggests, however, that these effects may not be about punishing dirty companies with adverse publicity, but about praising green leaders who attract business and investment. Amato and Amato found no negative effect on the stock values of the largest 500 US companies from being in the bottom quartile in *Newsweek's* ranking of 'The Greenest Big Companies in America'. In the 10 days after the ranking was announced, however, the praise impact on the share price of firms in the top quartile was very large. Their share price performance was more than twice as good (less than half as negative during a bear market) as the other 75 per cent of US firms.

Criminological theory needs something better than cynicism driven by piling up empirical studies about the limits of deterrence. Few citizens think deterrence has no role to play in the prevention of rape, theft or corporate crime. In failing to develop a theory of deterrence that takes fear and shame seriously in social control, criminologists have handed the deterrence debate to neoconservative maximalists.

This is sad because deterrence maximalists justify a greatly increased level of suffering for incarcerated people and their families. Maximalism increases crime by skewing state budgets towards prisons that mostly worsen the criminality of those sentenced to them (Nagin et al. 2009). Yet maximalists make more sense to the community than criminologists who say only that punishment deters little. One way of seeing the imperative to take deterrence seriously, even if minimally, is that deterrence underwrites the greatest historical accomplishment of the justice system. As discussed

in Chapter 3, Eisner (2003: 126) revealed sharply falling homicide rates between the sixteenth and seventeenth centuries across Europe—the period when some European states institutionalised courts to manage violence. The capability of courts in the eyes of citizens to deter and incapacitate violence was one reason they abandoned the private deterrence of blood feuds, thereby greatly reducing homicides (Pinker 2011).

A strategy of minimally sufficient punishment can increase the power of deterrence theory in crime prevention substantially if its empirical claims are more strongly verified by future research and if it can win the political debate against maximalists. It can increase the power of deterrence with a policy that involves the release of most prisoners even in societies with the lowest imprisonment rates. At the same time, it can tackle the unfreedom of overcrowded prison systems where viruses like HIV, tuberculosis, hepatitis, Covid-19 and violence itself proliferate, and human rights are compromised. Luckily, Ebola was contained in Africa before it caused terrible devastation inside the walls of overcrowded western prisons. Minimal sufficiency's claims are consistent with the limited existing evidence.

The aim in this chapter is to develop ideas towards a theory of minimally sufficient deterrence and to reflect on that evidence. To do that, the chapter discusses seven interwoven principles of a crime-prevention meta-strategy for minimally sufficient deterrence:

1. Escalate enforcement: Display intent to progressively escalate a responsive enforcement pyramid that involves progressive escalation of sanctions for wrongdoing and support for social responsibility.
2. Inexorability: Pursue inexorable consistency of detection of predatory crime. Communicate inexorable community commitment to stick with social support for those struggling with problems of lawbreaking until the problems are fixed.
3. Escalate social support: With repeated offending, increase social support. Even when there is escalation to a last resort of severe incapacitation, escalate social support further. Keep escalating social support until desistance is consolidated.
4. Sharpen the Sword of Damocles: Cultivate the perception that 'trouble hangs inexorably over my head; they want to support me to avert it'.

5. Dynamic concentration of deterrence: Focus deterrence on a line that should never be crossed after an announcement date. Then progressively lift that line, raising our expectations of socially responsible citizens and corporations.
6. Community engagement: Engage the community with offenders in widening restorative conversations that educate in the shamefulness of criminal predation for the many who participate in the conversations. Avert stigmatisation.
7. Modesty: Settle for the modest general deterrence delivered by this shamefulness and a minimal number of cases that escalate towards the peak of the enforcement pyramid.

This chapter first explains the idea of the responsive regulatory pyramid. It provides a scaffolding for these seven principles. Readers versed in the literature of restorative justice and responsive regulation may find much of the pyramid discussion familiar and can skip over it. Then the importance of inexorable response is explained, followed by the theory of dynamic deterrence and defiance and next how to constitute the shamefulness of the curriculum of crimes. The chapter concludes with an inductive conceptualisation of minimally sufficient deterrence.

The regulatory pyramid

The regulatory pyramid defines a meta-strategy of regulation, a strategy for how to sequence strategies (Braithwaite 2008). It is relevant to regulating crime by organisations or individuals. Figure 9.1 is an example of a regulatory pyramid elaborated in the next section. The presumptive strategy (a presumption that can be overridden) is to start at the base of the pyramid and escalate slowly. This is a strategy for keeping the power of deterrence sharp by making it rare to reach the pointy end of the pyramid. The rationale for keeping the Sword of Damocles sharp is discussed later.

Consistent with the evidence on what works with corporate crime enforcement from Schell-Busey et al.'s (2016) systematic review of 58 studies, this is a strategy that provides a wide mix of regulatory options before recurrently deterrent measures are reached that risk blunting the Sword of Damocles. At the bottom of the pyramid are restorative strategies that provide support to offenders and victims, and that meet needs and repair harms.

Figure 9.1 A responsive pyramid of minimally sufficient deterrence

Responsive regulatory theory says we should first look to the strengths of a lawbreaker and then seek to expand them. Mental health researchers led the way in showing that training in building on strengths improves quality of life and vocational and educational outcomes (Stanard 1999). When those outcomes are improved, recent econometric findings show more clearly than in the past that unemployment can be averted by vocational training and education, that this reduces crime and that wages for the poor can be increased, reducing crime by increasing the attractiveness

of legitimate work compared with illegitimate work (Chalfin and McCrary 2017: 33–35). This path to crime reduction is resource intensive, though less so than a massive prison system, and it is a benefit to the economy in contrast to the economic deadweight of prisons. The idea is to absorb weaknesses by expanding strengths. Put another way, regulators should not rush to law-enforcement solutions before considering a range of restorative approaches that can be delivery vehicles for capacity-building and for collective efficacy (as discussed further in Chapter 11). As some offenders see their strengths expand to levels not previously considered possible, societies must celebrate their innovation, publicise it and support its extension. With corporate enforcement, research grants and prizes for rolling out new approaches that take internal compliance systems up through new ceilings for that industry offer an illustration. An example is involving young black men in Minneapolis in celebration circles. Victims join with the offender's loved ones to celebrate the way an offender has repaired the harm, righted the wrong and turned their life around (Braithwaite 2002: 103).

As we move up the pyramid through a first to a second to a third restorative conference, conference participants are likely to decide to escalate to increasingly punitive interventions. The policy idea is to persuade participants that they should also keep escalating to new ideas and resources for providing support for the offender. The philosophy of restorative justice is to empower stakeholders to take advice from experts. But then stakeholders should make their own decisions contextually attuned to the circumstances of their offender. This includes knowledge of the programs the community of care can persuade the offender to complete and the needs of their victim and other stakeholders in their circle. A problem with this is that it does lead to a bricolage of community responses rather than one that maps mechanically from 'what works' criminology. This is a complex relational and community empowerment process in which the problem is in the centre of the circle, rather than a stigmatised individual being put in the centre. The content of the pyramid is not prescriptive. The use of terms such as 'escalated support', 'wide' and 'wider' support and 'escalated deterrence' without specifying escalation 'to what specific measures' is intentional.

With complex phenomena, it is best to follow not the most evidence-based strategy, but the best meta-strategy. The research question of which strategy works best is not as fertile as which meta-strategy works best. Rarely will the first strategy attempted work in a complex context that

differs from the conditions of any controlled evaluation trial. A good meta-strategy informs stakeholders of results from the 'what works' literature and presumptively tries the most supported strategy first and then the second most supported strategy (after the first strategy fails). That presumption can be overridden in light of particular circumstances. Clinicians, by analogy, try one therapy after another for a patient, informed by their knowledge of the outcomes of randomised controlled trials and their knowledge of particular cases, including what other medications patients are taking, their capabilities for surviving side effects, how strong their hearts are, and much more. As in restorative practice, clinicians can also decide what to recommend to catalyse community controls that will prevent spread of a contagion. In neither case does best practice involve a narrow focus on an individuated view of what works.

More detailed discussions of how to go about the process of deciding when and how to escalate up a responsive regulatory pyramid, and how to mobilise networked escalation as opposed to simple statist escalation, can be found in Braithwaite et al. (2007) and Braithwaite (2008, 2011). Intervention in complex phenomena like criminal careers should follow a trajectory that first assumes the answers are knowable and known. Therefore, evidence-based strategies can be applied (Braithwaite and D'Costa 2018: Ch. 12). When that fails, assume the challenge is knowable but unknown (and work to acquire at least some contextual qualitative understanding of the knowable). Then, if that repeatedly fails, assume one confronts a complex or chaotic phenomenon that is unknowable. In that situation, do not surrender to analysis paralysis; keep probing with new forms of social support that emerge from restorative conversations until a resonant response begins to produce positive change.

At the base of the responsive pyramid of sanctions are the most restorative, dialogue-based approaches we can craft for securing compliance with a just law. One reason for an approach that is deliberately responsive to complexity is that a particular law, or its interpretation, may be of doubtful justice, in which case we can expect dialogue mainly to be about the justice of the law. This is imperative for the law of freedom (Pettit 1997). If excessive force was used during arrest, or if racism was in play, we can expect dialogue about whether it is the defendant or police who have committed the greater crime. As we move up the sanctions pyramid, increasingly demanding interventions are involved. The idea of the pyramid is that our presumption should always be to start at the base, and

then escalate to somewhat punitive approaches only reluctantly and only when dialogue fails. Then we escalate to even more punitive approaches only when more modest sanctions fail.

Strategic use of the pyramid requires resistance to categorising problems into minor matters that should be dealt with at the base of the pyramid, more serious ones that should be in the middle and egregious crimes at the pyramid's peak. The presumptive preference, even for serious crimes, is to try dialogue first, overriding that presumption only if there are compelling reasons for so doing. There will be such reasons in exceptional cases: a violent first offender who vows to keep pursuing the victim to kill them may have to be locked up; a person who has never offended but attempts to blow themselves up in a subway may be killed by police who get a clear shot. The 2005 incident in which British police shot an innocent Brazilian man in a subway who was suspected of terrorist intent illustrates the justification for the responsive regulatory imperative always to consider, however quickly, the viability of interventions at lower levels of the pyramid.

As we move up the pyramid in response to repeated failures to elicit restorative reform and repair, in most cases, we eventually reach the point at which reform and repair are forthcoming, even if it is many years later. Whenever that point is reached, responsive regulation means that escalation is reversed; the regulator de-escalates down the pyramid. The pyramid is firm yet forgiving in its demands for compliance. Reform must be rewarded just as recalcitrant refusal to reform is ultimately punished.

A dramatic transformation of criminal law jurisprudence will be necessary if evidence supportive of responsive regulatory theory continues to mount. The imperative to de-escalate deterrence responsively when an offender rehabilitates means that every year a reformed person remains in prison is needless suffering. It is in addition a frittering away of society's scarce crime-control resources and a path to blunted power for deterrence.

If the empirical claims of responsive regulatory theory are right, this is also a missed opportunity to reduce crime and increase freedom by putting rewards for rejecting a life of crime alongside sanctions for embracing crime. In practical terms, when social support succeeds in helping prisoners serving long sentences to turn their backs on a life of crime, what is needed is a return to court for a hearing about the possibility of sentence shortening. The sentencing judge in such hearings should be obliged to

consider the views of victims who in turn listen to the opinions of parole professionals, the offender and the offender's family. From a responsive regulatory perspective, a criminal law that keeps people in prison until they have paid the proportionate penalty for their wrongdoing is folly and a threat to freedom. It is an indefensible policy in terms of a dynamic theory of deterrence. It can make deterrent sense only under a passive deterrence theory, especially a maximalist one. This is the passive theory that minimally sufficient deterrence seeks to render obsolete.

The deterrent superiority of the active deterrence of the pyramid is opposed to the passive deterrence of a fixed scale of consistently imposed penalties. This is elaborated in Braithwaite (2002: 73–136; 2008). The current state of the evidence for the superiority of integrating restorative justice and responsive regulation to secure compliance with the law and other outcomes important to domination reduction is traversed in Braithwaite (2021f).

Consistently proportionate punishment is justified by just deserts theories of equal punishment for equal wrongs. Equal punishment for equal wrongs, however, is a danger to freedom and justice. It privileges punitive equality for offenders, while riding roughshod over the justice claims of future victims of crime who suffer because of an inferior crime-prevention policy. In addition, some present victims may not want equal justice for equal wrongs to apply in their case. For example, they may prefer more reparation, less imprisonment and more support of offenders' family members who suffer to variable degrees because of an imprisoned breadwinner (Braithwaite 2002, 2003). In any event, what kind of equality is expressed by a logic of equal punishment for equal wrongs when some offenders are lucky enough not to be raped or bashed in prison and others do suffer these horrors? There is little equality of justice when some are trapped in prison-induced contagions of drug addiction and others are not; when some acquire HIV, hepatitis or tuberculosis in prison systems that are the best-known incubators of these contagions, and some do not. This is not to disagree that maximum sentences should be set on the basis of seriousness; it is to say that the right sentence is the minimally sufficient one.

Responsive regulation has had some influence on business regulation policies, but almost none on policing policies. This would not have surprised Edwin Sutherland (1983), who 70 years ago first demonstrated the propensities to tolerate forgiving approaches towards crime in the suites that are seldom evident towards crime in the streets.

Restorative justice provides stakeholders with professional advice on the rehabilitation and prevention options they might choose. The community of care can then be mobilised to monitor and enforce compliance with whatever is undertaken. This is an approach informed by values that define not only a just legal order, but also a caring civil society. These values are derived from the foundational republican value of freedom as nondomination (Braithwaite and Pettit 1990; Pettit 1997). Some who share these restorative values derive them from different foundations, including spiritual ones.

Ordering strategies in the pyramid is not just about putting less costly, less coercive, more respectful options lower in order to preserve freedom as nondomination and to save money. It is also that use of more dominating, less respectful forms of social control only after more dialogic forms have been tried first helps law enforcement to be seen as more legitimate. When regulation is seen as more legitimate and more procedurally fair, compliance with the law is more likely (Tyler 1990; Tyler and Huo 2002). Nagin and Telep (2021) question this, while agreeing with the perceptual claim. They argue that there is a dearth of evidence that actual procedural fairness increases compliance. Astute business regulators often set up legitimacy explicitly (Dekker and Breakey 2016). During a restorative dialogue over an offence, the inspector says there will be no penalty this time, but they hope the manager understands that if they return to find the company has slipped out of compliance again, they will have no choice but to refer this to the prosecution unit. If and when the manager explicitly agrees that this is a reasonable approach, a future prosecution will likely be viewed as fair. Parker's (2006) 'compliance trap' will then be less likely to trip up enforcement. Under this theory, therefore, privileging restorative justice at the base of the pyramid builds legitimacy and therefore prevents crime (Tyler et al. 2007).

There is also a rational choice account of why the pyramid works. System capacity crises result in pretences of consistent law enforcement when the reality is that punishment is spread thinly and weakly (Pontell 1978; Pontell et al. 2014). Unfortunately, this problem will be worst when lawbreaking is worst; criminal justice is a sprinkler system that fails when the fire gets hot. Hardened offenders learn that the odds of serious punishment are low for any infraction. Tools like tax audits that are supposed to be about deterrence can backfire by teaching tax cheats how much they can get

away with (Kinsey 1986).² The reluctance to escalate under the responsive pyramid model means that enforcement can be selective in a principled way. The display of the pyramid itself channels the rational actor down to the base of the pyramid. Noncompliance comes to be seen (accurately) as a slippery slope. In effect, the pyramid solves the system capacity problem by making punishment cheap. The pyramid says: ‘Unless you punish yourself for lawbreaking through an agreed action plan near the base of the pyramid, we will punish you more severely higher up the pyramid (and we stand ready to go as high as we have to).’ So, it is cheaper for the rational actor to self-punish (as by agreeing to payouts to victims or community service). Some Asian criminal justice systems, such as that of Japan, work this way much of the time, even for serious crimes such as rape, aggravated assault and murder, which are frequently resolved through compensation and remorseful apology rather than through prison time. Such reparative leniency does not cause crime to spin out of control in Japan (Ahmed et al. 2001). Once the pyramid succeeds in creating a world in which most punishment is self-punishment, there is no longer a crisis of capacity to deliver punishment when it is needed. One of the messages the pyramid provides to corporate criminals is that ‘if you violate repeatedly without reform, it is going to be cheap for us to hurt you (because you are going to help us hurt you)’ (Ayres and Braithwaite 1992: 44).

Paternoster and Simpson (1996) showed the limits of passive specific deterrence on intentions to commit corporate crime. When respondents held personal moral codes, these were more important for predicting compliance than were rational calculations of sanction threats (though the latter were important, too). Appeals to business ethics (for example, through restorative justice that exposes executives to the consequences for victims of a corporate crime) therefore may be a better first strategy than sanction threats (Parker 2004). It is best to succeed or fail with such ethical appeals first and then escalate to deterrence for the minority of contexts in which deterrence works better than ethical appeals. One of the psychological principles in play here is that when intrinsic motivation to comply with the law is intact, do not crowd out intrinsic motivation with extrinsic threats (Ayres and Braithwaite 1992: 49–50; Osterloh and Frey 2013; Frey 2017). This is another key to averting the deterrence

2 Mazzolini et al. (2017) found that, on balance, audits increased reported incomes by an average of 8 per cent, though audits that detected no extra tax liability reduced future reported incomes in the short term (see also Mendoza et al. 2017). In a meta-analysis, Dularif et al. (2019) found no effect of increased audits on tax evasion.

trap (Coffee 1981) and what Parker (2006) calls the compliance trap. Nine meta-analyses completed since responsive theory and behavioural economics picked up ‘crowding out’ and minimal sufficiency from developmental psychology have shown that there remains strong psychological evidence that crowding out does occur. They also show that both intrinsic and extrinsic motivations independently affect behaviour (Cerasoli et al. 2014).

According to responsive regulatory theory, what we want is a legal system in which citizens learn that responsiveness is the way our legal institutions work. Once they see the law as a responsive regulatory system, they know there will be a chance to argue about unjust laws or unjust enforcement (as opposed to being forced into a lower court production line or a plea bargain where such discussion receives short shrift). But defendants will also see that game-playing to avoid legal obligations inexorably produces escalation, as does failure to listen to arguments about the harms their actions are doing and what redress is required. The forces of law are listening, fair and therefore legitimate, but also ultimately might be viewed as invincible.

A paradox of the pyramid is that to the extent that we can guarantee a commitment to escalate if steps are not taken to prevent the recurrence of lawbreaking, escalation beyond the lower levels of the pyramid occurs in a low proportion of cases. This is the image of invincibility making self-regulation probable. Without commitment to escalation when reform fails to fix the problem, the system capacity crisis rebounds. A fundamental resource of responsive regulation is the belief of citizens in the inexorability of escalation if problems are not fixed.

Restorative justice works best with a spectre of punishment threatening in the background but never threatened in the foreground. When punishment is thrust into the foreground, even by implied threats, other-regarding deliberation is made difficult because offenders are pushed to deliberate in self-regarding ways—out of concern to protect themselves from punishment. This is not the way to engender empathy for the victim or internalisation of the values of the law or the values of restorative justice. The job of responsive regulators is to treat offenders as worthy of trust. When regulators do this, the law more often achieves its objectives (Braithwaite and Makkai 1994; Gangl et al. 2015; Haas et al. 2015). This ideal is one version of enculturating trust (in the foreground) while institutionalising distrust (in the background) through deterrence as a last resort.

Testing theories about dynamic interventions layered in a pyramid is more complex than testing the effects of passive policies like heavier sentences because the effects of sequences of interventions must be tested. How can a regulatory pyramid be tested when it involves an entire suite of sequenced dialogic, then deterrent and next incapacitating approaches? It has worked in raising an extra billion dollars in tax for each million spent on a program for multinational companies engaged in illegal profit shifting (for example, to tax havens; Braithwaite 2005b: 89–100). Evaluation in a tax compliance context requires, first, the creation of this whole pyramid of sequenced new policies for companies that have been paying no tax, and then observation of how much tax they pay after the new pyramid is put in place, as well as observing at what sequenced stage of the pyramid most tax payment starts to flow. The quality of information from the latter observations is instructive, yet low. This is because we do not know whether a compliance effect is the result of the last step up the pyramid or a combined effect of some subset of the whole sequence of escalations. A comparable evaluation challenge applies to problem-oriented policing as a meta-strategy. Randomising some police patrols to problem-oriented policing shows that problem-oriented policing works as a meta-strategy (Braga 2002; Weisburd et al. 2010; Hinkle et al. 2020), but it gives feedback of limited quality on which initiatives addressing which problems produced the result. Even so, evaluating meta-strategies is more important work for criminology than evaluating single crime-control strategies.

Braithwaite (2021f) has advanced an approach for how to think clearly about evidence in relation to dynamic theories of supports and sanctions. It argues that restorative justice and responsive regulation should each be viewed as constitutive of the other and both are dialogic forms of regulation. Restorative justice should be at the base of responsive regulatory pyramids and should be a practice of responsive listening. For both restorative justice and responsive regulation, motivational interviewing is an important evidence-based practice of active communal listening and active responsibility for problem-solving. Braithwaite (2021f) shows that both restorative justice and responsive regulation are meta-strategies; meta-strategies are strategies for deciding on and sequencing strategies of prevention.

Seven meta-analyses have all concluded that restorative justice has modest effectiveness in reducing crime. Braithwaite (2021f) argues that this is because it is a superior delivery vehicle for a great variety of strategies

that work. Responsive regulation works for the more banal reason that it involves trying one prevention strategy after another until the problem goes away. Van der Heijden (2020) completed a meta-review of a mix of qualitative and quantitative evaluations of responsive regulation. He found a positive effect of responsive regulation in eight settings, no effect in one, a negative effect in six and in nine other data settings observed effects that were qualified and context specific. Van der Heijden's accurate but underwhelming answer to the question of whether responsive regulation is effective was: 'It depends.'

When responsive regulation is found to be effective, this is because it covers the weaknesses of one preventive strategy with the strengths of many other strategies that are redundantly layered above it in the regulatory pyramid. Braithwaite (2021f) follows the lead of earlier work on why restorative justice and responsive regulation are vindicated by some empirical evidence. In that work, restorative justice and responsive regulation are seen, like community policing at hotspots, as street-level strategies that are problem-oriented and responsive to dialogue that builds collective efficacy. The collective efficacy theme is developed further in Chapter 11.

The empirical effectiveness of restorative justice and responsive regulation in preventing lawbreaking in part is seen as a result of both doctrines being grounded in a principle of street-level responsiveness. Braithwaite's (2021f) conclusion advances a view about how social science understanding has failed to cut through to why policing is wrongly seen as ineffective when policing that is fair and responsive can be very effective, why UN peacekeeping is seen as ineffective when it is even more effective in saving lives than the best hotspot policing, why rehabilitative programs are wrongly seen as ineffective, why regulation for corporate crime control is wrongly seen as ineffective and why foreign aid investments in development are wrongly seen as ineffective when they make very small contributions to a profoundly large and widespread problem of global poverty. These perceptions are in error, especially when interventions are responsive and grounded at the street level, because normal social science fails to see why good meta-strategies are more fundamental than learning to pick the most evidence-based strategy:

We can read the meta-analyses that suggest that problem-oriented policing works, that motivational interviewing works at the end of its iterated reframings of motivation, that positive deviance strategies

for improving village nutrition work, that a multidimensional mix of strategies works in controlling corporate crime (Schell-Busey et al. 2016) and that multidimensional UN peacebuilding works (Braithwaite and D'Costa 2018) as converging on a paradoxical insight. This is that, in a world of complexity, it is more possible to discover the meta-strategies that work best than it is to move single strategies from the realm of the knowable to the realm of the known. For example, the meta-strategy of 'search for positive deviance' may be more useful than learning what are the particular forms of positive deviance that worked to improve nutrition in particular villages. To use another example, it is easier to know that a vague, heterogeneous concept such as problem-oriented policing or motivational interviewing works than it is to know that it works because it fixes the street-lighting at hot-spots or discovers some specific motivation for losing weight. And this is a methodologically impressive paradoxical finding because it is harder to muster the statistical power to show the efficacy of heterogeneous than homogeneous interventions. (Braithwaite and D'Costa 2018: 553)

Inexorability of supports and sanctions

Inexorability has three elements:

- prioritising increased consistency of detection above tougher punishment
- always taking serious crimes seriously with a continuum of restorative responses to every detected serious crime; avoiding 'do-nothing' responses
- escalating the seriousness of response to a second, third and fourth offence; sticking with the problem until it goes away.

Prioritise detection

The inexorability piece of the theory of minimal sufficiency builds on the evidence from the deterrence literature that the perceived and actual severity of punishment are rarely good predictors of compliance with the law, while the perceived and actual certainty of detection are often useful predictors (Blumstein 2011; Robinson 2011; Friesen 2012; Nagin 2013). One reason for this is that detection mobilises not only formal punishment but also informal disapproval, which is a more powerful driver

of compliance with the law (Braithwaite 1989). Theoretically, this is not just about the evolution of cooperation (Axelrod 1984) or the evolution of compliance when noncompliance is visible to a punisher. The newer theoretical insight is that it is also about indirect reciprocity through fear of reputational loss even without repeated encounters with the same people (Berger 2011; Nowak 2012; Braithwaite and Hong 2015; Hong 2016). Criminologists therefore tend to read the deterrence literature as showing that ‘detection deterrence’ and ‘disapproval deterrence’, both specific and general, are more powerful than deterrence by severe formal sentences. Minimally sufficient deterrence is based on this view that ‘detection deterrence’, indirect reciprocity and ‘disapproval deterrence’ are indeed thin reeds that can be combined to be more powerful than deterrence by severe state punishment. This conclusion runs contrary to the Nobel Prize-winning predictions of Gary Becker (1968), which enjoy little empirical support according to this reading of the evidence.

Always respond

Inexorability is absent in contemporary urban justice. Enforcement swamping and system capacity overload mean that young people picked up as minor first offenders learn that they do not receive significant punishments even if they are prosecuted. This is also likely to happen with their second, third or fourth minor offences during their teenage years. When the system does finally decide to hit youth offenders hard because someone decides they have ‘had enough chances’, offenders wonder ‘why now?’. Legitimacy is a casualty of this policing strategy for muddling through system capacity crises. Tough punishment seems to repeat offenders to have unfairly come out of the blue, when they got away with worse in the past and when they see friends get away with even worse. Because this seems and is arbitrary, it is by definition a threat to freedom (Pettit 1997, 2014) and has shallow legitimacy in their eyes. These dynamics are quite comparable with Parker’s (2006) compliance trap for corporate crime. The next section considers an alternative response approach to first, second, third, fourth and fifth offences.

Escalate responses

Critics might say the trouble with inexorability is that it is hard to reconcile with minimal sufficiency of punishment. Punishing everyone detected indeed would be maximal net widening rather than minimal

sufficiency. The challenge of averting net widening is to craft a minimally sufficient response for a minor first offence. Police, teachers or parents who observe children hitting each other do well to pause rather than walk by. They must insist that the children stop fighting; they do best to say something non-stigmatising like, 'You guys are better than that', and then walk on after assuring themselves that fighting has ceased. This is a better way of taking violence seriously than looking the other way. It is more than 'nattering' as one walks by without stopping the violence (Patterson and Bank 1989), but less than net widening, which creates a recording of alleged wrongdoing and a formal decision on what to do about it.

Restorative theory can inform an inexorability that averts the perception of an arbitrary punishment lottery. The evidence is strong that restorative justice buttresses the legitimacy of the justice system (Tyler et al. 2007; Sherman 2014; Barnes et al. 2015; Miller and Hefner 2015). Prosecution is not the way to go with a first-offending child arrested for a petty offence. Nor is turning a blind eye. Wang and Weatherburn (2019) found that police cautions for minor youth offenders resulted in lower reoffending than arrest and referral to juvenile court. A restorative police caution with a degree of ritual seriousness is an option. Police can respond to a shopkeeper holding a child who has stolen something by ensuring the child returns the stolen property, taking the shopkeeper's contact details and then taking the child home to ask their parents or guardians what they intend to do. Or the child can be held at the police station until their parents arrive to take them home following a restorative caution.

The restorative caution gives the child and parent space to come up with the suggestion that they will visit the shopkeeper together to apologise, perhaps even baking a cake or bringing some flowers. Traditionalists see such idiosyncratic gestures of apology as strange elements to take seriously in criminal justice policy, yet that is the essence of trusting the community rather than the police with averting an offender's reaction that 'nothing happened, so breaking the law is no big deal'. Police tell the parents they expect a text advising what has been done to apologise. The police say they may check the shopkeeper is satisfied. In other words, most of the work of social disapproval is delegated away from the police. A reason for this as one approach to taking every crime seriously is the evidence that censure by families and closest friends is more likely to be a reintegrative form of shaming, while censure by criminal justice officials is more likely to be stigmatising (Ahmed et al. 2001: 157–76).

So, what to do about the teenager's second minor offence? The minimally sufficient deterrence suggestion is a restorative justice conference that the victim is invited to attend. The child's loved ones would be expected to sit in the circle for a serious family ritual involving parents, grandparents, siblings, perhaps aunts, uncles and a sports coach or a teacher trusted and nominated by the child. Communicating this expectation is important because a concern is to ensure that overburdened mothers do not shoulder all the burdens of social support. Wider circles of participation also enhance the effectiveness of restorative justice (Braithwaite 2002: 50–51, 55, 74, 252–65). This result is also evident in Wilson et al.'s (2017) meta-analysis finding that teen courts, impact panels and reparative boards are ineffective forms of what some loosely call 'restorative justice', but which in fact are programs that are thin on collective participation and collective efficacy.

Unlike a criminal trial that assembles people who can inflict maximum damage on those on the other side of the case, the restorative justice conference assembles people who can offer maximum support to their own side, be it the victim's or the offender's. At a meeting of two communities of care, the communication of disapproval comes from those personally affected by the crime but, more importantly, also from those who most love the offender. Nathan Harris's evidence from restorative justice conferences is that only disapproval communicated by people the offender most loves is effective in inducing remorse (Harris 2001: 157–76). People who are well liked but not loved are not potent at inducing remorse. Nor are the police. Albert Bandura (2016) made the point that self-censure for cruel conduct is switched off when others are stripped of the quality of being humanly important to us. Indeed, Bandura suggests that not only restorativeness but also responsiveness are 'capabilities' in Sen and Nussbaum's sense of capabilities being fundamental to freedom. Responsiveness is learnt through the loving social capital of families: '[I]nterpersonal experiences during the formative years, in which people experience joys and suffer pain together, create the foundation for empathic responsiveness to the plight of others (Bandura 1986)' (Bandura 1999: 200). Humanising those whom one might want to hurt or punish is fundamental to preventing moral disengagement (Bandura 2016). For Bandura, the psychological evidence shows that self-efficacy is a key to strong, free individuals resisting pressure from delinquent peers to dehumanise others so that crime and other predations might be inflicted on them (for example, Bandura et al. 1996).

While an informal police caution for a first offence is a minimalist response in terms of taking the crime seriously, a restorative justice conference for a second offence escalates to a longer family and community ritual with a trained facilitator and a wider circle of participation by people concerned about the child. Such a conference becomes a focused way of supporting children. Are they struggling in school? If so, what support can the conference mobilise? Are they struggling in their relationships? Are their friends leading them into trouble? Is there support from other friends who can steer them clear of such trouble? If there are problems with alcohol, drugs or anger management, proactive support may be needed. In this world of social support, every child leans on a 'youth support circle'. This is a restoratively elaborated version of parent-teacher conferences in schools that meet every year with every child aged over 12, with their extended family and with well-networked elders until the child is helped to get his or her first job or into university (Braithwaite 2001). The youth support circle is designed to reduce stigmatisation of crime by being universal; children who never do anything wrong have the circles. In that world of a better-funded, more communal welfare state, this conference for a second offence has no extra cost because it would be integrated into routine youth support conferences for building human capital, affecting only the timing of a conference that might normally be annual.

What about a third criminal offence? A longer restorative conference with a wider circle of participants is needed, usually with a follow-up conference to celebrate completion of an agreement. That would be more onerous than the conference for the second offence. More importantly, the next conference would see an escalation of social support for the child compared with the first conference. A child welfare worker could attend. The expertise a trained social worker brings would include knowledge of the range of options available in the town for rehabilitation of the young offender. The social worker should also have knowledge of the principles of risk-need-responsivity in evidence-based selection of rehabilitation options (Andrews and Bonta 1998, 2010), sound knowledge of the 'what works' literature of criminology and a good clinical capacity for responsiveness to the complexities of the specific case. In a restorative justice conference, it is not the job of the expert to dictate to a family (Pennell and Burford 2000). Restorative justice works by delivering stronger implementation of conference agreements enforced by the parties themselves than courts can achieve with police enforcement. This was the biggest effect size in the Canadian Department of Justice meta-analysis of restorative justice by Latimer et al. (2001). The effects

of completion of restorative justice agreements were stronger than the statistically significant effect of restorative justice on reduced reoffending compared with control group members.

We can reconcile these results by understanding that if a restorative justice conference and a court both send a child to a counterproductive program, restorative justice will do more damage. The child will be more likely to complete the counterproductive program when it was agreed to by the family and other conference stakeholders than when the same outcome is ordered by the court. Restorative justice does greater harm than court when the circle agrees to counterproductive measures, and greater good than court when it agrees to effective measures. The reason is that restorative justice is a superior delivery vehicle for the completion of rehabilitation programs.

The idea is to strengthen this comparative advantage of restorative conferences by investing in experts who speak up when the family considers sending the child to a scared-straight program. The expert points to the evidence that scared straight is counterproductive (Weisburd et al. 2017: 428). It follows from this that a good way to reanalyse a meta-analysis such as that by Lipsey (2009) would be to assess whether effect sizes can be increased by the combination of highly effective interventions such as social cognitive programs with restorative justice as their delivery vehicle. Put more provocatively, it is not useful to compare effect sizes for restorative justice with those for other programs because restorative justice should be conceived of as a way of delivering multiple strategies—a meta-strategy (Braithwaite 2021f). It makes more sense to compare restorative justice with court as an alternative delivery vehicle for diverse correctional options, as in Strang et al. (2013). Likewise, the end of this chapter argues that it makes more sense to compare deferred corporate prosecutions with corporate convictions than to compare deferred prosecutions with some other approach to corporate reform. The bigger insights might come from teasing out which specific combinations of programs and delivery vehicles have positive and negative synergies, as is done in the smart business regulation literature (Gunningham and Grabosky 1998).

A conference for a fourth offence might allow the family to mobilise rehabilitative options from further afield or expensive options that are rationed. Critics might query why such an expansion of the quantum of social support would make a difference given that in Lipsey's (2009: 141) meta-analysis of youth justice programs, providing more hours of services,

surprisingly, did not increase the effect sizes of interventions. Restorative justice programs were the big exception to this result; hours of restorative service provision strongly increased the already statistically significant effect size of restorative justice in reducing reoffending. Within the 'restorative justice' category of programs, those that included a mediation component, as opposed to simple restitution, also had an effect size more than one-third higher (Lipsey 2009: 142).

A conference for a fourth offence might also send the conference option to court for approval (or modification) by a judge, perhaps as a deferred prosecution. A meta-analysis from seven British studies led by Joanna Shapland concluded that restorative justice conferences have benefits that average eight times as much as their costs (Strang et al. 2013: 44–46). This result is a reason not to consider costly escalation to court until a fourth offence. Yet isn't escalation to something less cost-effective inept at any stage? Actually, there is a relevant complexity in the evidence that should leave us open to this. While Strang et al.'s systematic review found that court is clearly less effective in preventing crime than restorative justice, it also suggested that a combination of court and conference could be more effective than either on its own. More data are needed to assess whether this is robust. The quantitative transitional justice literature finds that war crime prosecutions, truth (and reconciliation) commissions and amnesties all have limited or contextual explanatory power on their own. However, when all three are used together, combining the punitive and the restorative, countries experience strong reductions in human rights abuses. This is particularly true when the truth commission's engagement with civil society is wide and deep and when amnesties are qualified rather than blanket (Olsen et al. 2010; Dancy et al. 2013; see also Sikkink 2011: 184–87). The combined cost of a restorative justice conference that then reports to court might also be less than the sum of its parts if the integration can be designed to streamline court processing. This is essentially how the most comprehensive youth justice conferencing program in the world operates, in New Zealand (Johnstone 2013).

At a fourth conference (for a fifth offence), when a young person and a victim are on the precipice of deeper trouble, escalated interventionist expectations can be assumed by the community of care. For example, in a 2014 interview, I was told of a teenager in an Irish Republican Army area of Belfast who had repeatedly assaulted his mother. The restorative conference was conducted by Community Restorative Justice Northern Ireland. One part of the community restorative justice agreement, which

had many parts, was that four community members agreed to respond immediately to calls for help from the mother and participated in training on how they could respond. These were not civil servants living far away, arriving the next morning. They lived around the corner and committed to respond within minutes, 24 hours a day. This is a good example of how restorative justice can expand to a wider, more timely and more proximate web of social support and social control, while still providing a softer web of collective efficacy than court enforcement to protect the mother by locking up her child.

By this point in our inexorability narrative, deterrence maximalists are aghast that this is a fifth criminal offence with no formally punitive response presumed. The offender has had ‘five free hits’: the police restorative caution, followed by four restorative justice conferences and a first deferred prosecution that might involve a court appearance—all ‘doing nothing’ for punitive deterrence. Perhaps there will have been six ‘free hits’ if the first restorative caution was preceded by an informal warning on the run. All I advocate for these mostly disadvantaged fifth-offender children is that we give them five chances in the same way that we do for corporate criminals, as explained in the final part of this chapter.

Contrary to maximalist fears, offenders perceive restorative justice conferences not as ‘doing nothing’, but as gruelling experiences meeting their victims in the presence of their loved ones (Umbreit and Coates 1992; Schiff 1999). Deterrence maximalists are wrong to see imprisonment as the only kind of perceptually tough response. Perceptually, *The Process is the Punishment*, as in the title of Malcolm Feeley’s (1979) book, especially when the process is designed with a ritual seriousness that is emotionally demanding.

If offending persists, repair of harm does not occur, but restorative justice achieves far from nothing because we resiliently stick with the problem. We refuse to take the easy path of putting the offender into a punishment production line that casts them out of our sight. The resilience to stick with the problem is accomplished by empowering those who most love the offender and the victim to stick with the problem with state and civil society support, especially from people who are passionate about the social movement for restorative justice. That passion is an ingredient that cannot be achieved without patient work to build a movement for restorative justice. Restorative justice people actually ask us to tarry with the problem in the centre of the circle. Many who had not been

restorative justice people had this experience of seriously tarrying for the first time with the problem of white supremacy and black oppression in their emotional engagement with the killing of George Floyd in 2020 by the weight of a policeman's knee. In June 2020, after National Rugby League games in Australia, the two majority-white teams would mingle in a circle arm-in-arm, led by black-minority players to 'take a knee' together to remember Floyd and reflect in silence on how they were going to be actively responsible for confronting racism in Australia. Many around the world in that month would tarry for eight minutes and 46 seconds to represent the duration of the policeman's knee on Floyd's neck that stopped him breathing.

Building a social movement for restorative justice is a slow-food paradigm shift for injustice in which the passion to struggle against domination is accomplished by asking us to pause with the story of a single victim and perpetrator, or two. And then for an active politics of scaling up that reflection and commitment even across to so many rugby league teams in Australia. We scale up what Chantal Mouffe calls agonistic pluralism (Chapter 12) because of the power of the movement, the power of the story and the power of love for a suffering human being that makes us human. Without the narrative and the global anti-racism movement, we never renarrate the politics of denial of racism.

Dynamic deterrence and defiance

The Sword of Damocles is an ancient metaphor popularised by the Roman senator and republican philosopher Cicero (1877: 185–86). He based it on the story of a Sicilian king who hung a sword attached by a horsehair above the head of a courtier called Damocles, who envied the king. The ruler wanted to illustrate the insecurity of being king. Today, the Sword of Damocles generally refers to any ever-present peril hanging over a person. The existence of ever-present peril is an important element of minimally sufficient deterrence.

Preserving the sword

At a child or young adult's appearance for a sixth offence in a criminal career, the court might signal that a sword hangs over them. This is best done not as a threat. The power of the sword, according to Cicero, is

not that it falls or is threatened; its power is that it hangs. The regulatory literature shows that the best signalling is for the judge to say at the outset of the court hearing for a sixth offence that its objective is to support the family and save them from having their child taken away. Perhaps only later than a sixth offence and only after a very serious crime would the judge ask whether the offender would think it reasonable that they be incarcerated to protect the community were they to commit another serious offence. Note the use of motivational interviewing techniques here, which are empirically established as effective and that avert threat-making (Rubak et al. 2005; Lundahl et al. 2010). The objective at that later trial is to open the mind of the offender to the reasonableness of the community protecting itself with a custodial sentence.

The idea at the next trial for the next serious offence would be to remind the offender that in their last appearance they said themselves that a custodial sentence would be reasonable if an offence of this seriousness recurred. The judge would then concede the offender's point of view but mobilise social support for one last chance to stay in the community, while making it clear that next time they were likely to agree with the incapacitation recourse that the offender themselves had concluded was reasonable in these circumstances. Lorana Bartels (2009) showed that a suspended sentence can be an effective Sword of Damocles here. She found that suspended sentences in Tasmania resulted in low reconviction rates compared with executed sentences.

At every stage, the minimally sufficient approach requires that the offender be led to see a new escalation of social support provided in response to a new transgression, but also a set of punitive options with a long prison term at the peak of the pyramid. Community service orders, fines, electronically monitored home detention, orders to a violent husband to transfer bank accounts to his wife so she has the financial capacity to leave him and a diverse variety of other options that are found lower in the pyramid are available as alternatives to prison. The sheer diversity of community gifts of support conveys a message of care when it includes, for example, the Royal Society for the Protection of Cruelty to Animals program in Australia that guarantees care to the pets of domestic violence victims who would otherwise stay in abusive relationships to care for those pets. The escalation of support as a life careens into deeper trouble is a way of increasing the legitimacy of more severe sanctions as a last resort when escalation to them does occur. It is also a strategy for combating the widespread perception by criminal offenders in many societies that the

system lets you get away with it for years and then one day out of the blue locks you up. The proposal is inexorable both in escalating support and in the way it signals a move to escalating deterrence.

Why reserve court appearances until after a fourth or fifth officially detected offence? Why reserve serious sanctions for later still in a criminal career? One reason is the evidence from randomised controlled trials of restorative justice in Canberra led by Lawrence Sherman and Heather Strang. Offenders randomly assigned to restorative justice had greater fear of future criminal enforcement after restorative justice conferences than offenders randomly assigned to criminal prosecution had after their trial (Braithwaite 2002: 119–22). Offenders emerged from restorative conferences more fearful that they would be rearrested if they offended again, more fearful of family and friends finding out and more fearful of a future conference, compared with those assigned to court (Sherman and Strang 1997; Sherman et al. 1998). Minimally sufficient deterrence favours restorative justice for multiple early offences in a criminal career because these data show that restorative justice sharpens the Sword of Damocles, sharpening deterrence. Criminal trials blunt the sword hanging over the offender. After the courtroom sword is brought down, its mystique is lost. The criminal trial in current judicial practice blunts deterrence because in most non-serious cases the offender is surprised at how easily they get off as the court struggles with system capacity overload.

The minimally sufficient deterrence idea is to hold the trial in reserve until it is time to take the case seriously by projecting a clear trajectory of escalation to an ever-bigger Sword of Damocles that is being averted by more and more support. Among other objectives, this support is intended to make that sharpening of deterrence appear ever more just. Other kinds of criminological evidence support a Sword of Damocles effect (Sherman 1992, 2011), including Dunford's (1990) finding that a warrant for arrest deterred domestic violence substantially better than either actual arrest or nonarrest. Later, this chapter considers the possibility that deferred individual and corporate prosecutions may deter better than completed prosecutions. The theoretical perspective of minimal sufficiency is that warrants for arrest have great attractions over actual arrest and that deferred prosecutions are more powerful tools than actual prosecutions. These are problem-solving tools that can enable support to play a larger role than sanctions. Concluding that deferred prosecutions are in principle powerful tools is not to deny that their widespread use in corporate criminal law has often approximated doing nothing in matters of a seriousness that called

for doing quite a lot (Eisinger 2017). Nevertheless, we will see in the final sections of this chapter that, empirically, restorative corporate justice can also sharpen deterrence while conviction of the corporation can blunt it, actually causing a criminal corporation's stock market value to rise.

Dynamic deterrence that accounts for defiance

Responsive regulatory theory argues that the passive deterrence thinking of the law and economics tradition, as in Gary Becker's (1993) Nobel Prize-winning work, has limited value. The reason is that real-world deterrence unfolds dynamically. Dynamic deterrence moves through sequences of threats; passive deterrence is static, involving levels of threat that are constant across time. International relations theorists have been more dynamically sophisticated than criminologists and economists of deterrence. They do not assume that, even though the United States has a bigger deterrent arsenal than the rest of the world's militaries, it works for America to say to another country: 'Do what we say or else!' There is evidence aplenty, from countries as close to the United States as Cuba, that threats are as likely to induce defiance as compliance. This is accepted even by conservative writers like Michael Rubin who oppose dialogue with 'rogue states'. Rubin (2014: 4) nevertheless conceived of Cuba, North Korea, Iran, Iraq and Libya as 'backlash states' that were 'defiant'. Former US defence secretary William Cohen tweaked this definition of rogue states to conceive of them as regimes 'immune to traditional deterrence' (Rubin 2014: 4). While demands for compliance backed by passive deterrence work poorly in international affairs, when the United States dynamically escalates its deterrent power towards a weaker country, as it did during the Cuban Missile Crisis, it can get a deterrent result (dismantled Cuban missiles). Dynamic escalation of deterrence in international affairs is a dangerous game because little Cuba might mobilise powerful friends to dynamic escalation of their deterrent capabilities in response. Little Serbia did manage to dynamically escalate catastrophic deterrence by triggering the escalation to World War I after the assassination of Archduke Franz Ferdinand by Serbian terrorists.

Psychologists of learning approach the way punishments work as dynamic learning sequences that are beyond the writ of static deterrence models. They demonstrated psychological reactance to threats (Brehm and Brehm 1981). Defiance is the more elegant term that Sherman (1992) deployed to describe this phenomenon.

Figure 9.2 Theory of the effect of coercion on compliance as the net result of a capitulation effect and a defiant resistance effect

Sources: Based loosely on experiments summarised by Brehm and Brehm (1981) and subsequent research (Rains 2013).

A paradox of the pyramid is that by being able to escalate to tough responses at its peak, more of the regulatory action can be driven down to its deliberative base. Yet punishment, according to responsive regulatory theory, simultaneously increases deterrence and defiance. Figure 9.2 is a way of summarising the implications of more than 50 experiments on defiance originally conducted by Brehm and Brehm (1981) and their colleagues, and many more since (for example, Rains 2013). At low levels of punishment, defiance usually exceeds deterrence. Figure 9.2 expresses this as the resistance effect exceeding the capitulation effect at lower levels of coercion. The dashed line is the net compliance effect represented as a sum of the resistance score and the capitulation score. Only when punishment bites very deeply at the peak of the pyramid, resulting in many giving up on resistance, does deterrence exceed defiance.

Yet escalation only as far as the lower levels of the pyramid can elicit compliance when that first step up the ladder is seen as a signal of willingness to redeem regulators' promises to keep climbing until the problem is fixed. Put another way, the first escalation becomes a wake-up call that convinces offenders they are heading towards deterrence. Social

support initiatives also help by signalling that paths away from deterrence are available. A perception of the dynamic inexorability of the pyramid does most deterrence work, not the passive general deterrent.

Redundancy in the design of the pyramid saves the day when defiance effects initially exceed deterrence effects. The redundancy idea is that all regulatory tools have deep dangers of counterproductivity. Therefore, one must deploy a mix of regulatory tools with heavy representation of dialogue and social support in the mix. The best way to deploy the mix is dynamically, so that, in sequence, the strengths of one tool have a chance to cover the weaknesses of another. For example, the pyramid in Figure 9.1 is about the strengths of one form of restorative justice covering the weaknesses of other forms of restorative justice, strengths of deterrence covering weaknesses of restorative justice, strengths of incapacitation covering weaknesses of deterrence and strengths of strong social support covering weaknesses of limp social support. Hipple et al. (2014) found that restorative conferences that avoided defiance by being strong on restorativeness and strong on procedural justice were more effective at reducing reoffending. The risk of defiance exceeding deterrence is one reason the peak of the pyramid should always be threatening in the background, but not directly threatened in the foreground. Making threats increases defiance, turning the resistance curve in Figure 9.2 more steeply downward. How, then, can lawmakers and business regulators be threatening in the background without making threats? One way is by being transparent that the pyramid is the new policy. Law enforcers must be the change they want to see. They achieve this by communicating openly with society about the policy design of the pyramid. By citizens being invited to be partners in the democratic design of different regulatory pyramids for responding to different crime problems, citizens also come to learn about the inexorability of escalation until the law redeems its promises. The ideal is to communicate the inexorability of deterrence in this way rather than by making threats in specific cases.

Dynamic deterrence as a remedy to enforcement swamping

A dependable, inexorable peak to the pyramid is a particular way of thinking about what Mark Kleiman (2009) calls dynamic concentration of deterrence—often called (in a misleadingly static way) focused

deterrence.³ For responsive regulation, the dynamic concentration of deterrence potency is at the rarely used peak of the pyramid. Kleiman, like David Kennedy (2009), reached a kindred conclusion to responsive regulatory theory about the superiority of dynamic over passive deterrence through contemplating how to respond to enforcement swamping as a challenge for thinly resourced policing agencies.

Kleiman's dynamic concentration theory shows why abandoning random targeting in favour of strategic concentration of targeting can work as long as monitoring works. In the simple case of scarce resources enabling the targeting of only one of two regulated actors, an erroneous intuition is that 'concentrating on Al would allow Bob to run wild'. If Al is promised certain punishment, a rational Al will comply if the compliance costs are less than the penalties. 'Then Bob, seeing that Al has complied, will himself comply; otherwise, Bob knows that he would certainly be punished. So, giving priority to Al actually increases pressure on Bob.' In this we see the dynamic elements of the strategy. Kleiman (2009: 54) shows that this initial insight holds for a variety of conditions such as promising certain punishment of the second mover rather than the first, and larger numbers of players. Dynamic concentration helps a little punishment go a long way.

Tax authorities likewise have learned how to respond to enforcement swamping when rich people, trusts and companies stampede into illegal tax shelters. This is to announce that while the tax authority lacks the resources to enforce the law against all who herd into shelters, they can prosecute the first risk-taker to jump into a shelter after the date of their announcement of intent to attack particular shelters in the courts. This can be extremely effective in ending cascades of risky tax cheating by high-wealth individuals and corporations (Braithwaite 2005b). Braithwaite (2012) discovered the same dynamic concentration in the wisdom of

3 The danger of describing the theory behind innovations like Operation Ceasefire as focused deterrence is that it will be understood as a static policy of identifying the highest-risk group for targeting. Even the principal authors of the strategy, who clearly understand its dynamic qualities, often describe a static deterrence targeting strategy, complemented by short breakouts into discussing its dynamic aspects (Kennedy et al. 2017). The most common mistake regulators make concerning responsive regulation is to understand it as a static policy of triaging the highest-risk groups for targeting with tougher deterrence. The point of reframing deterrence is to push criminologists away from such static ways of thinking. Minimally sufficient deterrence commends restorative justice as an alternative to prison even in the highest-risk circumstances such as creeping genocide, actual genocide or murders that risk further revenge killings (for example, Braithwaite and Gohar 2014).

generals who faced the biggest enforcement swamping challenge in the world at that time: small numbers of UN peacekeepers facing Africa's worst war in Democratic Republic of Congo.

As usual, practitioners here were ahead of theory. Tax officials were ahead of us, as were those generals in Congo and that Texas Ranger on the screen in our youth. The Ranger faces a lynch mob with one bullet in his gun. He turns them away with the promise that 'the first to step forward dies'. Kleiman (2009: 49–67) elegantly theorises why the dynamic concentration of deterrence by the Texas Ranger works. Systematic reviews of dynamic concentration found consistent effectiveness across studies and a medium-sized statistically significant crime reduction effect overall (Braga and Weisburd 2012, 2014; Weisburd et al. 2017; Braga et al. 2018, 2019). The intuition that concentrating deterrence on Max will allow Mary to run wild turns out to be wrong in terms of rational-choice theory (Kleiman 2009: 49–67) and empirically wrong according to Braga et al. (2013: 315), who found that with dynamically concentrated deterrence, 'vicariously treated gangs were deterred by the treatment experiences of their rivals and allies'. Dynamic focusing on the peak of a pyramid is just one way of concentrating limited enforcement resources that delivers dynamism to both specific and general deterrence.

Dynamically raising the bar serious offenders must jump

Boston's Operation Ceasefire is criminology's *locus classicus* of the dynamic concentration of deterrence in showing how an inner-city justice system overwhelmed by gang violence reduced homicide (Kennedy 2009). It was also 'focused deterrence' in that it did not attempt to deter all crimes perpetrated by gang members, only their gun crimes. My hypothesis is that the passively focused features of Operation Ceasefire may have some value, but its dynamic concentration of deterrence is more innovative and more germinal. It follows that rerunning and updating of Braga and Weisburd's (2012) encouraging meta-analysis are needed to compare those interventions that were simply focused and passive in their deterrence with those in which the intervention delivered a dynamic concentration of deterrence. The dynamic concentration aspect of Operation Ceasefire involved the Texas Ranger trope described above. Police sat down in meetings with gang leaders and members to let them know, in effect, that we know that you know that we have insufficient capacity to go

after all of you for all your offences. But we do have the capacity to go after all the offences and all the parole and probation breaches of all the members of the next gang to use a firearm in a crime.⁴ This means that instead of concentrating deterrence on the worst offenders, deterrence was dynamically concentrated on the first offender to use a firearm after the announcement date. The theory of the intervention was that all gangs would self-regulate gun carrying and use to avoid being the first gang to be targeted or the second gang to be targeted after the first. The ethnographic side of the evidence on the formidable desistence of these gangs from gun use seems to support this hope (Kennedy 2009). For proportionality theorists, it is a weakness of the program that it diverts resources from prosecuting the most serious offences to what might be minor parole violations after a gang uses guns. This critique also applies to responsive regulation. It has been eloquently advanced by Karen Yeung (2004).

Operation Ceasefire was in tension with the minimally sufficient deterrence model in two ways, however. First, the approach was thin on restorative justice and social support as approaches that strengthen a deterrence strategy into which restorative justice and social support are integrated. There was certainly dialogue with the gang members involved, and pathways out of the gang were discussed and even provided for some; but from the perspective of minimally sufficient deterrence, it was too narrowly oriented towards pulling levers to focus deterrence as more 'swift and certain' for a strategically targeted subgroup. There are reasons to suspect that in some of these programs this swift and certain deterrence may have been communicated with Trumpian threat. Such threats risk being counterproductive according to minimal sufficiency theory, defiance theory and the theory of motivational interviewing.

There was insufficient attention in Operation Ceasefire to a dynamic approach to support. Leaders of the innovation protest that this is not true (Kennedy et al. 2017). If moderately violent societies like the United States are to learn how to manage their hotspots better from experience with peacekeeping in extremely violent societies like Congo, we might

4 Papachristos and Kirk (2015: 536) describe the moderator of what they found to be the effective Chicago program opening before that with: 'This isn't a trick. Everyone gets to go home tonight. So relax a bit. We're here to talk to you about one thing: gun violence. No tricks. Just some straight talk, and an offer to help.' And, indeed, unlike some of these initiatives that put all the emphasis on deterrence, the encounter moves recurrently to: 'If you want help, it's here for you.' Service providers in the room explain what help they can offer, including health, mental health, housing, drug treatment, education and employment service providers (Papachristos and Kirk 2015: 537).

be able to build a consensus for the dynamic concentration of support to become more prominent. This policy lesson has been better learned in international peacekeeping and peacebuilding than in national policing to control organised violence (Braithwaite and D'Costa 2018: Ch. 3). That might be one reason for the conclusion in Chapter 11 that UN police are more effective than domestic police in saving lives (Hultman et al. 2013). The lesson is that desistance should not only cause a lifting of punishment; the strategy should also maximally concentrate rewards and supports at the moment of desistance. The rewards are not only tangible matters of vocational training and job placement, but also rituals of pride at celebration conferences in which loved ones eulogise peacemaking and rehearse redemption scripts (Maruna 2001).

Project HOPE is a drug court program in the focused deterrence tradition that initially seemed to have promising pyramidal features of escalated responses targeted at hard cases. HOPE stood for Hawai'i Opportunity Probation with Enforcement. It has been adopted in dozens of US locations with 'Honest' replacing 'Hawai'i', yet with mostly limited investment in creating job or other 'opportunities'. Intervention escalated as drug users went off the rails. Yet it may be the program that ran off the rails. Much of the rhetoric of its practitioners was maximalist, oriented to 'swift and certain' deterrence. This happened when the evidence is not supportive of criminal justice swiftness (Pratt and Turanovic 2018), though swiftness of a supportive parental firm hand in childrearing is important. Hawken and Kleiman (2009) entitled their evaluation *Managing Drug Involved Probationers with Swift and Certain Sanctions: Evaluating Hawaii's HOPE*. Duriez et al. (2014) raise the concern that the ideology driving the diffusion of Project HOPE has emphasised its 'swift and certain' character, ignoring other positive features such as motivational interviewing training for officers in the program—something for which there is a strong evidence base (Rubak et al. 2005; Lundahl et al. 2010), which is why motivational interviewing has become central to restorative and responsive justice (Braithwaite 2011). The findings of the literature reviews can be characterised as landing somewhere between HOPE showing great promise and being discouraging of HOPE (Lattimore et al. 2016).⁵ As with Operation Ceasefire, the systematic reviews should

5 Lattimore et al. (2016: 1103) describe their HOPE program's four-site evaluation of a 'program that emphasizes close monitoring; frequent drug testing; and swift, certain, and fair (SCF) sanctioning. It also reserves scarce treatment resources for those most in need.' There is not much escalation of support in that description, nor any dynamic distinctiveness of the deterrence strategy to transcend the limits of static deterrence.

be reanalysed after some on-the-ground engagement with what each specific program does. They could be coded qualitatively or quantitatively according to four variables: how much deterrence is involved (HOPE's 'E' = 'Enforcement'), how much social support is involved (HOPE's 'O' = 'Opportunity'), how dynamically concentrated is the deterrence and how dynamically concentrated is the support. Meta-analyses might contribute more to science if they were more theoretically focused on what they evaluate and less focused on heterogeneous puzzles like HOPE that are in essence brands.

The second tension between minimally sufficient deterrence and Operation Ceasefire is that an approach that says your gang will be targeted intensively only if it uses guns challenges the inexorability principle. In an enforcement swamping crisis, however, we must confront the reality that priorities must be set that start where it is most important. In Democratic Republic of Congo, that priority was mass rape atrocities in which sometimes hundreds of women and children were raped, murdered or enslaved into mines (Braithwaite 2012). At least one peacekeeping commander was effective in reducing this seemingly impossible enforcement swamping crisis during the first decade of this century, according to our Congo fieldwork. He convinced assembled militant leaders that the next militant group to commit a mass rape atrocity would be the group on which peacekeepers would focus all their military capabilities to bring perpetrators to trial. In fieldwork trips to Congo during the past decade, I reproached the head of the UN Organization Stabilization Mission in the Democratic Republic of the Congo, the military commander, the deputy commander, the general in charge of the relevant region and the US ambassador for failing to implement this strategy against Colonel Cheka of Mai Mai Cheka, allegedly the worst perpetrator of mass rape atrocities, and some others like him. Within a year of the appearance of publications that discussed this (for example, Braithwaite 2012), the United Nations announced a policy shift in the direction of the dynamic concentration of deterrence, though this had nothing to do with the publications. This did improve the security of the people of Congo after 2014, particularly through the surrender of the M23 armed group that in 2012 captured Goma—a strategic city with a population of 1 million that is now liberated. In 2017, after almost three years of sustained military pressure to force surrender to face trial, Cheka turned himself in. That is no more than suggestive qualitative evidence

for the dynamic concentration of deterrence from the least likely case (Eckstein 1975) of probably the world's most extreme and persistent enforcement swamping crisis of recent decades.

Gun violence was obviously an outstanding target for Operation Ceasefire. It produced a wonderful result in reducing shootings and homicides by more than 30 per cent—a result that continued to be supported in more recent work on the dynamic concentration of deterrence on US gang violence (Braga et al. 2013; Kennedy et al. 2017). The strategy, however, fails the inexorability test of minimally sufficient deterrence because inexorability happens only for the offence in focus (usually gun crime). Prioritising the greatest harm is desirable and might not deeply threaten inexorability as long as there is a strategy to move on to clean up one kind of gang crime after another, to move down to the B-list of gang harms and then a C-list, after the A-list of gun homicide harms has been tamed.⁶ Then a strategy like Operation Ceasefire perhaps in the long run could pass the inexorability test.

Similarly, with the enforcement swamping crises with tax shelters that the United States and Australia faced in recent decades, Australia learned that it is possible gradually to raise the bar on tax compliance obligations. This was achieved by targeting the 10 worst tax compliers each year—a different 10 each year because last year's terrible 10 are no longer so terrible this year:

When the judgment is made that there is a culture of tax cheating in a particular market segment, the industry norm revealed in the multivariate analysis is still used to target those furthest below the norm for audit and other compliance tools. But more of them can be targeted than in other industries. And when they are caught out by the audit, the bar they are required to reach before they are released from targeted surveillance can be raised a little higher than the industry norm. As a result of the worst 10 compliers in the industry moving from way below the old norm to above it, the norm of course is moved upwards. Then in the next year, a new set of the worst 10 in the industry is moved up above that higher norm. This raises the bar again. We can in this way keep raising the bar with problem industries until they are paying their fair share. (Braithwaite 2005b: 160)

6 At times, practitioners speak of A-lists and B-lists in static terms, by, for example, arguing that police go after an A-list of the most serious offenders for automatic prosecution, putting only the B-list into an Operation Ceasefire or a deferred prosecution program.

Stampedes of the wealthy into tax shelters do recede, as they did in Australia in the late 1970s and again around 2000 (Braithwaite 2005b). Cascades of open-air drug markets taking over great cities—even the stairs of the New York Public Library—also end, and that contributed to the downward cascade of homicide in Manhattan (Zimring 2011). Consider a brief list of accomplishments in reversing catastrophic cascades. The ozone hole was substantially closed even though it seemed unstoppable until the Montreal Protocol started to reverse the cascade into chemical use that was widening it (Kuttippurath and Nair 2017). Resources were provided to developing countries to comply with the Montreal Protocol after 1987, and there were diplomatic shots across their bows as well, particularly by US embassies (Braithwaite and Drahos 2000: 261–67). Today, we see substantial beginnings of reverse cascades from coal power to solar and wind energy. President John F. Kennedy predicted a cascade to 15–25 nuclear powers by the 1970s, yet half a century beyond the 1970s only the United States, Russia, China, the United Kingdom, France, Israel, India, Pakistan and North Korea have nuclear arsenals. This is thanks to civilising forces in international civil society that won a 2017 Nobel Peace Prize and dogged regulatory inspection in places like Iraq under the nuclear nonproliferation regime (Braithwaite and Drahos 2000: 318). Interminable civil wars in the places where the worst wars have cascaded for longest, such as Congo, will also one day reverse to cascades of peace (Braithwaite and D’Costa 2018; Walter et al. 2020).

The trend in regulatory theory is to seek ever greater sophistication in risk assessment and risk management as the main game of how best to cope with the seemingly impossible challenges of regulatory enforcement swamping (Black and Baldwin 2010). Though this is not totally wrong, there are also risks in shifting a high proportion of regulatory resources into deskbound risk analysis positions in a regulator’s central office and away from street-level inspection. We may learn most from worst cases like Cambodia after its multiple cascades of genocidal violence beyond the 1970s. In the 1990s, and particularly since 1998, downward cascades of violence began to spread in Cambodia. Broadhurst et al. (2015, 2018) described this as a dynamic civilising process (Elias 1969; Pinker 2011). Broadhurst and his colleagues document that local police and UN peacekeepers did useful things to help trigger the reverse cascades. They show with Cambodia, as a least-likely case (Eckstein 1975), that it was not so much that Cambodian police were geniuses of risk analysis. Rather, they did something that Malcolm Sparrow (2000) describes simply as ‘pick important problems; fix them’.

Cambodia was more a matter of a return to long-run momentum towards the civilising processes that citizens crave and governments pursue when they seek legitimacy from citizens and from the international community. This was combined with police and peacekeepers helping with an A-list of violence problems that they helped clean up, eventually moving on to B and C-lists. A-list criminality included robbery, homicide and kidnapping, with cattle theft being high on the B-list because this can be financially devastating in rural societies. Local police became quite popular, according to Broadhurst and his colleagues, and gradually moved away from putting bullets in the heads of desperadoes and towards peacetime community policing.

We can learn from local priests, mayors and elders in Rwanda who acted like Texas Rangers without even a single bullet in the midst of Rwanda's cascade of genocide. These leaders stood their ground—in most cases, stopping the genocide from spreading to their community through their emotional dominance in insisting that their church, their leaders, would not stand for this in their village (Klusemann 2012). Other brave priests who tried this were hacked down. Together, long-run civilising processes helped by dynamic concentrations of sanctions and support, combined with gradually raising the bar on what kinds of violence are intolerable, eventually can pacify even genocidal intent in Cambodia, Rwanda or Timor-Leste, as documented in the previous chapter. It can also close an ozone hole and end the slave trade to the Americas. All this can be accomplished without filling prisons.⁷ Or so I hypothesise.

Conversations across the curriculum of crimes

Restorative justice principles are useful to a minimally sufficient deterrence strategy because defiance (Sherman 1992, 2011) is a critical risk. Defiance is reduced when communities of care do most of the work. Nathan Harris found that perceived informal disapproval from those most loved inside

7 Rwanda did fill its prisons with 126,500 people charged with participation in its 1995 genocide. Many were children who were then raped in prison and died from HIV/AIDS. Many others were innocents forced to participate after seeing their family members hacked to death for refusing. Defendants were executed on the judgements of second-year law students. In the end, that tiny, poor society did not have the capacity to deliver justice to 126,500 defendants for such serious crimes. Most were released to face traditional *gacaca*, which much of the time was somewhat restorative in approach (Clark 2010).

restorative justice conferences, not from criminal justice actors, does the work of persuading offenders that their crime is shameful, persuading them to remorse and repair of the harm to victims and to their own family (Harris 2001: 157–76). In restorative justice, there is no need for anyone to invoke the concept of shame, nor for anyone directly to shame offenders. Loved ones discussing how concerned they are about the consequences of the crime, the suffering of victims and what the family can do to help repair the harm is the way to elicit remorse without defiance. Motivational interviewing of these loved ones can draw this out.

Conversely, there is evidence that stigmatisation (as opposed to reintegrative shaming) increases crime in criminal justice processes (Ahmed et al. 2001; Braithwaite et al. 2006; Tyler et al. 2007; Braithwaite 2020c) and in business regulation (Makkai and Braithwaite 1994b; Harris 2017). Stigmatised offenders are treated as bad people who have done bad things, while reintegratively shamed offenders are treated as essentially good people who have done a bad thing. Stigmatised offenders are cast out from the community of the law-abiding without paying attention to reintegration rituals that might have drawn them back into the law-abiding community. Aversion of stigmatisation is critical to an effective package of minimally sufficient deterrence.

The theory of reintegrative shaming concludes that shame is important to crime control and to problem-solving (Leach and Cidam 2015; Spruit et al. 2016; Braithwaite 2020c). Societies in which rape is not shameful have a lot of rape. Societies in which feminist politics communicates the shamefulness of rape and domestic violence can enjoy steeply reduced rates of these crimes, as Pinker's (2011: 196–201) analyses of declining rates of rape and domestic violence and the growing shamefulness of these crimes in certain western societies suggest. Broadhurst et al. (2015: 310–13) likewise diagnose repeated surveys in Myanmar since 1996 to show declining domestic violence, growing disapproval of wife beating and growing public awareness campaigns about why it is wrong. Feminist politics is just one kind of engagement around the shamefulness of certain crimes.

One of the virtues of deliberative forms of justice such as restorative justice is that they increase the active participation of citizens in their democracy through the judicial branch of governance, through children's participation in antibullying programs in schools and through the involvement of environmental activists and fishers in the regulation of environmental crimes. Restorative justice therefore has a role to play in

educating citizens in the curriculum of crimes and why they are shameful, through their participation in restorative conversations about the crimes of their classmates, their neighbours, their family members and themselves.

Existing criminal justice institutions by contrast are overly professionalised. One consequence of this is they provide no space for democratic deliberation with the young about why crimes that affect people are wrong and what should be done about them. Democratic citizens can sit in the public gallery for criminal trials, but few do; and if they try to participate in the conversation about the rights and wrongs of the matter from the gallery, they are silenced.

Penal populism that increases punitiveness is certainly a risk within contemporary criminal law jurisprudence (Lacey 2008). Advocacy for minimally sufficient deterrence, however, is advocacy of quite a radical transformation of these dysfunctional institutions. Ordinary people are more punitive than the courts when they read accounts of cases and sentences in the media. When they read about the rich complexity of the same case as the judge hears it, they recommend sentences similar to those of the judge. The more information they have, the less punitive they are (Doob and Roberts 1983, 1988). In Warner et al.'s (2017) survey of jurors in Australian cases, they were twice as likely as the judges in their cases to recommend a noncustodial sentence. When citizens have the chance to engage even more directly with offenders and the complexity of their circumstances in a restorative conference, their punitiveness and vengefulness reduce even further (Strang et al. 2013: 40–42). This helps explain why restorative justice conferences produce less-punitive outcomes on average than do traditional criminal justice processes (Braithwaite 2002: 146–48).

In sum, the restorative justice component of minimally sufficient deterrence calms defiance,⁸ helps educate offenders and the entire community about the shameful of crime and the curriculum of crimes, while laying foundations for minimally sufficient punishment that can defeat penal populism's maximalist politics. A utopian world can be imagined in which each year 1 per cent of the population takes some responsibility for an offence in a restorative justice conference conducted

8 This also means that restorative justice and the escalation of deliberative supports help provide an exit from what Christine Parker (2006) calls the compliance trap, as discussed earlier and later in this book.

by the criminal justice system, a school, a university or a workplace. If 10 supporters of victims or offenders attend each conference, conversations about the curriculum of crimes would ripple across 10 per cent of the population each year. Because humans are storytelling animals, we learn the shamefulness of the curriculum of crimes through participating in, and retelling, stories of which we are a part. This retelling can do most of the work of constituting the curriculum of crimes, especially when newer crimes such as profit-shifting by multinational corporations become transparent. Consciences are formed by operation of 'the criminal law as a moral eye-opener' (Andenaes 1974: 116–17), especially when shamefulness is suppressed through a politics of domination with crimes such as torture and gendered violence.

Minimal sufficiency of general deterrence

The preceding section was partly about the general deterrence that arises from citizens talking with one another about why crimes like rape or torture are wrong. Reintegrative shaming theory advances the idea that general deterrence by means of the internalisation of shame (anticipated self-shaming rather than shame sanctions) combined with a path out of shame (Leach and Cidam 2015; Spruit et al. 2016; Braithwaite 2020c) are more important than deterrence by sentences handed down by courts (Ahmed et al. 2001). It is also about restorative justice's political strategy for community support for a less punitive justice system. Satisfaction with the justice, with the respect for victim rights and with the effectiveness of restorative justice in crime prevention is high (almost always over 80 per cent, and normally over 90 per cent) for citizens who sit in on restorative justice conferences (Braithwaite 2002: 45–71; Wilson et al. 2017). Part of the practical politics of driving punishment down to minimally sufficient deterrence is convincing politicians who see restorative justice as a soft option to sit in on a circle and chat afterwards with participants. This is important because democratic politics is the key constraint on whether and how judges or police can move to minimally sufficient deterrence.

The literature on the consequences of police strikes (Andenaes 1974) has long persuaded criminologists that crime spikes when deterrence is taken off the table. The contention of minimally sufficient deterrence is that courts will have limited work to do in delivering minimally sufficient general deterrence if citizens are empowered conversationally about the shamefulness of the curriculum of crimes (Braithwaite 1989: 77–79). Courts

must ensure through some form of incapacitation that the community is protected from the modest number of people who are a severe danger to the community. The hypothesis is that minimally sufficient punishment will be provided by general deterrence resulting from imprisonment in such incapacitation cases. These are combined with other general deterrence effects when repeated failures of social support and moderate deterrence escalate to more severe deterrence towards the peak of a pyramid.

So long as deterrence does not fall to zero, increasing average prison terms does not have much effect in reducing crime (Nagin 2013; Chalfin and McCrary 2017). It seems unlikely that any society will face crime risks from insufficient passive general deterrence if it takes seriously shame management and education about the curriculum of crimes and if it puts in place a credible peak as a last resort in its pyramid of dynamic deterrence. We cannot completely do without passive general deterrence, but a minimally sufficient quantum of it delivered by the model proposed here may be enough to achieve the limited work general deterrence can do.

My proposals can be accomplished only incrementally; learning through monitoring is important to reveal any explosion of crime driven by a deficit in passive general deterrence (Braithwaite and Pettit 1990: 140–55; Dorf and Sabel 1998). If and when empirical evidence suggested this was happening, incremental movement could be halted and adjusted to bolster passive general deterrence. The prediction, however, is that as societies such as the United States and Russia, with imprisonment rates of more than 600 per 100,000, reduce their passive general deterrence towards that of societies such as India, Indonesia and Japan, with imprisonment rates in the thirties to forties per 100,000, passive deterrence deficits will not cause crime waves. This view is encouraged by cross-national comparisons of crime that show that low-incarceration societies, many of which are in Asia, often have low crime rates.

Restorative justice strengthens deterrence of corporate crime?

This chapter has explained how responsive regulatory escalation makes it rational for corporations to punish themselves at lower levels of the pyramid to avert more punitive measures higher up the pyramid. I have already explained how restorative justice near the base of a regulatory pyramid can reduce what Parker (2006) calls the compliance trap.

It attempts this through the deep listening and motivational interviewing restorative justice requires and by minimising the degree to which extrinsic threats crowd out intrinsic motivation to comply with the law. Chapter 6 argued for softening the defiance and legal cynicism of the compliance trap by leveraging soft targets who do not benefit from the crime to secure compliance. Abandoning hierarchies of accountability in favour of circles of accountability was another important part of this argument. All this in turn is enabled by greatly pluralising the separation of private and public powers. Chapter 7 argued for social capital strategies of CHIME for securing compliance dialogically most of the time. Inspectors and auditors explaining initial forbearance from deterrence means deterrence is more likely to be perceived as fair when it does come. At the same time, clear signalling that escalation to deterrence and incapacitation is inexorable deters gaming that undermines the regulatory order for everyone. This includes those who perceive others to be playing the regulator like a fiddle. In this chapter, explanation of the power of restorative and responsive justice has moved to how rationing corporate deterrence can avert a blunted power of deterrent escalation through overuse, instead allowing restorative justice to sharpen deterrence, as it has been shown by some limited empirical evidence to do.

But there is more than all of these points involved in why restorative justice embedded in wisely implemented responsive regulation can strengthen the deterrence of corporate crime. The system capacity crisis is more profound with corporate crime than with crime in the streets because business regulators receive less funding than the police—even though corporate crime investigations are more complex and expensive (Coffee 2020). Acquittals for corporate crime are more common partly because not all lines of inquiry receive the investigative work they need. What restorative justice delivers here at the base of the regulatory pyramid is an invitation to a deferred prosecution or a prosecution forgone in favour of an enforceable undertaking if the defendant adopts various self-enforcement measures.

Spalding (2015) perceives US sentencing guidelines as encouraging restorative justice in deferred corporate prosecution agreements. We will see that one problem is that deferred prosecutions and enforceable undertakings became too routine in failing to honour the restorative principle of earned redemption. Coffee (2020: 147) has discussed the imperative for deferred prosecutions to be ‘truly earned’. Laying a charge, but formally deferring prosecution, is the quintessence of the Sword of

Damocles. The US Department of Justice's policy has stated in various ways at various times that it will not proceed from deferred to actual prosecution if the corporate defendant:

- voluntarily discloses the corporate offence and responsible individual offenders
- disgorges illicit gains
- makes a 'credible and authentic commitment to remedy wrongdoing and promptly self-report future violations of law' (Baer 2021: 351)
- 'invests significant resources in compliance-related activities' (Baer 2021: 351)
- cooperates fully with the government's investigation
- repairs the conditions that caused or promoted the alleged offence (Arlen and Kahan 2017).

Declining to move to an actual prosecution from formal notice of a deferred prosecution is supposedly the practice only if the corporate defendant makes a fair fist of these conditions. In the Corporate Enforcement Program under the *Foreign Corrupt Practices Act*, self-disclosing companies also avert Stanley Sporkin's disliked innovation of a third-party monitor appointed by the Justice Department. These benefits of self-enforcing justice are denied in theory to 'recidivists', though recidivism is a loosely defined concept. In practice, recidivists frequently receive deferred prosecution agreements as separate subsidiaries of a mega-corporation that is too big to nail. On an extremely wide front, they get non-prosecution.

There is experimental evidence (Bigoni et al. 2012) that cartel formation can be deterred when what the Justice Department's policy calls 'leniency' is offered to the first reporting party combined with high fines for parties that fail to voluntarily report offences. The problem in practice is whether firms that do not meet the Justice Department's requirements for corporate leniency are prosecuted. Generally, the answer is they are not. Often there is no reality to the inexorability of escalation to the tough peak of the enforcement pyramid. Hence, critics reasonably opine that leniency is a terrible name for the Justice Department's policy.⁹ Its beautiful theory

9 We might welcome Baer's (2016: 1113) plea for parsimony and clarity in corporate criminal prosecutions rather than 'leniency'. Of course, I often will prefer 'restorative' corporate prosecutions.

of principled contingency of leniency has become an ugly practice of consistent leniency towards corporate criminals, particularly during the Trump administration (Garrett 2020: 116).

Defenders of the policy say that even under President Trump a benefit of the leniency program has been that it has motivated massive growth of corporate investment in internal compliance programs. There is little doubt this is true; corporate compliance staff and the compliance consultancy market have accelerated almost exponentially (Baer 2021). Defenders of leniency include a high proportion of corporate law practitioners who point out that it has always been true that a tiny proportion of corporate wrongdoing is detected by the state and put into the hands of prosecutors. Most of the corporate wrongdoing that is detected, stopped, punished by action against responsible executives and then leads to repair for victims and prevention of recurrence is a result of detection accomplished by corporate compliance staff. Alternatively, detection is by internal whistleblowers who refuse to report to outsiders but will report noncompliance to compliance staff or a board ethics/audit committee. This is also undoubtedly true, though it is also a reason Sporkin's innovation of a government-mandated third-party monitor with a desk inside the firm should not have had its momentum reversed as much as it has.

In the 1980s and 1990s, a network of reformers of which I was a small part in Australia, associated with the Australian Competition and Consumer Commission and some other regulators, became heavily involved in seeding the corporate compliance market and corporate compliance professionalism in Australia. It seeded corporate compliance courses in law, business and accounting schools, formed professional associations for compliance and successfully urged Australian regulators to give more emphasis to this in their regulatory strategies. While we still take some pride in that networking, the critiques of our work 30 years on in Australia seem correct to me, and much the same as the US critiques.

Baer (2021) argues that the corporate compliance market is a market in lemons. Soltes (2021) captures the voices of critics of the compliance market who say it is one that 'finds loopholes to circumvent obstacles' or is a market in 'schooling executives in cover-up rather than compliance'. In Australian aged care regulation, Braithwaite et al. (2007) had been advocates of the growth of an aged care compliance consultancy market that assisted nursing homes to innovate to achieve 'continuous

improvement' in the Australian regulatory regime that required quality of care beyond compliance with static standards. Empirically, Braithwaite et al. (2007) found this innovation corroded to become regulatory ritualism. Ritualism is another way of describing Baer's market in lemons or Parker and Gilad's (2011) corporate compliance 'window-dressing'. The compliance consultants would arrive with ritualistic checklists that would not provide long-term improvements in the quality of care. For example, some short-term, unsustained reform that aged care residents would notice as an improvement might be put in place. Then a slipshod piece of survey research would be administered to residents or their families that would show the residents subjectively perceived a short-term improvement in their quality of care. This was compliance on the cheap, tick-the-box, short-term and short on integrity. Inspectors would find the same problem a year or two later. It was the antithesis of corporate restorative justice. Rather, it was a crude transaction of non-enforcement in exchange for compliance ritualism that saved criminal corporations losses they would have sustained under tougher enforcement. This reveals a problem with narrowly economic theories of why deferred prosecutions are good policy in circumstances of enforcement swamping and scarce prosecutorial resources (as in Leone et al. 2021).

This is a general problem across a diversity of domains of corporate compliance. One impressive regulatory initiative in the United States has been regulators orchestrating sanctions and incentives to motivate corporations to put in place internal whistleblower policies. Decades on, Soltes (2020) found that regulatory mandates that organisations have whistleblower hotlines (for example, under the *Sarbanes–Oxley Act*) resulted in regulatory ritualism. Soltes made field inquiries regarding alleged misconduct to the whistleblower hotlines of 250 firms. Receipt back of clear and specific answers and action on the inquiries were disappointingly infrequent. Worse, Soltes (2020: 429) found that one-fifth of firms had 'impediments that hinder reporting'—for example, disconnected phone lines, email bounce-backs, incorrect directions to a website. These hotlines not only fail to have a no-wrong-door policy, they also have a shunt to a wilfully closed door. In these cases, beautiful paper policies on whistleblowing were an ugly practice of deadend accountability. This is why the *qui tam* policies discussed in Chapter 6 that motivate insiders to blow the whistle on cover-ups of corporate crime

by paying them a percentage of the fine are a vital inexorability-inducing reform. It is why that reform has proved effective in the face of such weaknesses of corporate compliance reforms.

When Baer (2021) coins the idea of a compliance market in lemons to describe these developments, she tracks Nobel Laureate George Akerlof (1970) on how the used car market works. Buyers struggle to discern the difference between lemons and good used cars. Owners of cars in good shape drop out of the market because buyers do not trust their claims about the car and will not pay what it is worth. Thus, the market unravels to a point of market failure in which only crooks who tout lemons are willing to play. It is not an exaggeration to say that some markets in corporate compliance services are kindred criminalised markets in smoke and mirrors. Modern regulators monitoring compliance systems suffer an information asymmetry similar to buyers of used cars, especially the large number of regulators who have moved away from street-level inspections that kick the corporate tyres, in favour of desk auditing or algorithmic compliance. This is why Baer (2021) insightfully diagnoses the compliance market as having unravelled from an idealistic practice of late twentieth-century reformers from within corporations, regulators and civil society with an incipient restorative justice imaginary to a market in compliance lemons.

There are things that can be done to repair this market by a combination of market and regulatory means. For example, when an Australian aged care compliance consultant pushes away regulatory enforcement with a flawed compliance innovation that is poorly evaluated, the regulator should mandate the lodging of the evaluation report on the *My Aged Care* website. Then family and residents in the home can complain to the regulator that as soon as the regulator went away, the so-called reform amounted to nothing and long-term compliance worsened. Even more importantly in market terms, when compliance monitoring reports are put up on public websites, we know from programs like registered self-certification of software programs for tax compliance (Braithwaite 2005b: 87–89) that competitors in the compliance market go to the regulators to tell them that particular evaluations of compliance improvement are flawed. More often the reputable compliance practitioner goes to the firm that has hired the slipshod compliance professional to suggest that the job could be done more professionally by replacing their existing compliance consultant. When they do this in the market for compliance services, the firm that has bought a compliance lemon might see a risk that this competitor in the compliance consultancy market could alert

the regulator to the corrupted compliance work. Unfortunately, most compliance evaluation reports in contemporary regulatory practice are not transparently posted on accessible websites. A case can be made for internal compliance reports to be confidential when they relate to matters that are unknown to regulators. The argument is that if they were made public, firms would have less incentive to invest in internal detection and remediation. But there is no such argument in cases where a compliance report is produced pursuant to an enforcement action over noncompliance known to the regulator.

Baer (2021) points out that the evidence is encouraging that voluntary self-reporting of foreign bribery has become relatively common in the United States. In the theoretical terms of this chapter, the offer of deferred prosecution or non-prosecution has increased the inexorability of detection. After a corporation's counsel voluntarily discloses to a regulator or prosecutor that bribes have been paid, the prosecutor must decide if this has been a prompt and complete disclosure.

This, in turn, requires a fair amount of verification. The prosecutor might ask the corporation's counsel for a list of employees who have already been interviewed, for documents that have already been searched, and so on. The government might subpoena documents independently to corroborate the information it has received or conduct its own interviews of relevant witnesses. Verification is costly (Baer 2021: 359).

The more time and energy governments invest in testing the corporation's claims, the less valuable deferred prosecutions become as a solution to enforcement swamping and the system capacity crisis (Pontell 1978). Eventually, the regulator or prosecutor therefore decides to truncate their verification of voluntary reporting.

Public interest in the deferred prosecution bargain can erode because of adverse selection akin to the used car market Akerlof described. Ethical or remorseful corporate leaders can decide that disclosure is not worth the trouble and is against the interests of shareholders. Unethical executives continue to game the voluntary compliance system with compliance ritualism or compliance corrupted into sophisticated cover-up.

Major punitive prosecutions of firms that corrupt voluntary compliance professionalism and of firms that cover up are one important remedy. Another is requiring firms to evaluate the effectiveness of their compliance programs—something evident in only 55 of 255 US deferred prosecution and non-prosecution agreements (Garrett 2014: Ch. 3).

It remains a fundamental point that if that remedy is inexorably executed, restorative justice for genuinely contrite and reforming corporations could save more prosecutorial resources for those corporations that do game the law. This more strategic use of finite deterrence resources can make deterrence more potent. If restorative justice can genuinely deliver corporate compliance in those cases where contrition and reparation are volunteered because the justice is more genuinely restorative, the sword of corporate deterrence can be sharpened by putting more resources into the cases that need the most deterrence, the cases in which corporations are most dangerous because they are the least contrite and the most disposed to cover up.

Most fundamentally of all, Chapter 6 argued that it is wrong to evaluate corporate deterrence by how well it deters offenders. Corporate deterrence works best when it deters soft targets who are third parties with the power to prevent corporate offences rather than when it seeks to deter offenders before or after their offence. This was illustrated in Chapter 6 by Mitchell's (1994a, 1994b) work on the almost total ineffectiveness of the regime designed to deter shipping companies criminally responsible for oil spills, compared with the 98 per cent effectiveness of the regime that deterred the firms that insured and 'classified' those ships. It was illustrated by the Australian Taxation Office's campaign against profit-shifting by multinationals that raised \$1 billion in extra tax for every \$1 million spent on the program by targeting major accounting firms as gatekeepers rather than the offending firms themselves (Braithwaite 2005b: 89–100). Deterrence repeatedly is shown to work well by moving the corporate targeting away from a tough-nut corporate deterrence target and on to a soft but strategic gatekeeper or other third-party target with the capacity to prevent. This is because in corporate life the capacity to prevent is overdetermined and not primarily in the hands of individual offenders. As Cumming et al. (2021) and Dyck et al. (2010) argue, it takes a whole village to detect financial crime, which includes auditors, compliance and ethics staff, analysts, short sellers and institutional investors. Perhaps it is wrong to call this targeting deterrence.

That critique may be of no concern when the objective here is not to defend deterrence in any narrowly conceived way. On the other hand, when the International Maritime Organization as regulator says to the insurer that it will not be authorised to issue ship insurance unless it gets ships to do certain things, the regulator withholding the licence to insure

ships from an insurer is a form of deterrence. Likewise, it is a kind of deterrence when the insurer says to the ship owner that if you fail to do this, your insurance is void.

Another important thing to say about how corporate deterrence works is that it mostly works *before* sentences are imposed when cases do go to court. Chapter 6 discussed Waldman's (1978) and Fisse and Braithwaite's (1983) early research on this. It showed that the costliest things convicted corporations do in response to a prosecution are done prior to trial to improve their case for corporate responsibility presented at the trial. This is also why the stock market impact of state enforcement tends to come with the announcement of the prosecution or the investigation or even rumours of investigation of irregularities (Carberry et al. 2018), while 'the public corporation's stock price usually goes up on the announcement of the sanction' (Coffee 2020: 66). Karpoff et al. (2008b) support this pattern of a reputational effect of investigation rather than a sentencing effect in SEC cases in the United States. They further found that

the expected loss in the present value of future cash flows due to lower sales and higher contracting costs—is over 7.5 times the sum of all penalties imposed through the legal and regulatory system. (See also Karpoff 2012)

This reality opens the door to creative future use of restorative justice in research and development on deferred prosecutions and enforceable undertakings (Parker 2004). This is a particular instance of the more general point made about the Sword of Damocles in this chapter. The deferred prosecution process can be designed to sharpen deterrence, while imposition of a criminal sentence by a court blunts it with surprising frequency.

A final important point is that with occupational health and safety inspections and low-level fines, the evidence points to an almost total absence of general deterrence, but formidable specific deterrence (Scholz and Gray 1990). This may be because of the range of semi-formal and informal inspection compliance levers discussed in the context of Table 6.1. These levers can be readily enabled by backing inspectors with restorative circles that demand repair and future prevention, mostly of the self-incapacitation kind discussed in the next chapter. Likewise, regulators can be enabled to underwrite them by prosecutions (and indeed enabled to underwrite restorative justice). More precisely, the specific deterrence

and other preventive effects of inspections in Table 6.1 may be better underwritten by an integration of restorative justice and punitive prosecutorial justice than by either alone.

Too big to fail; too big to nail?

In the earliest days of restorative innovations in Australian Competition and Consumer Commission enforcement (Parker 2004), some of us at the ACCC did have the view that the victim compensation and other remedies agreed to in enforceable undertakings often had higher costs for the firm than would have been obtained by a prosecution. This was not because we saw ourselves as oppressive in our negotiation of these undertakings; we saw ourselves as tough negotiators who were firm but fair. And we as regulators were the initiators of Australian law reforms that sought more accountability in enforceable undertakings to ensure the terms were neither captured nor oppressive, and satisfied the rule of law. In early cases like the consumer frauds in remote Queensland Aboriginal communities by global insurance corporations, we believed the outcomes were tougher than judges would have imposed in a prosecution because top management, CEOs or board chairs of some of these companies had become genuinely ashamed of what their company had done. That was partly because there were some criminal convictions of individual insurance company executives as well. Top management had not started with remorse; that was an accomplishment of restorative elements of the process when CEOs sat with Aboriginal elders in ‘yarning circles’ (Parker 2004). Defendants often started with ridiculous techniques of neutralisation of responsibility that accused the accusers of oppressive enforcement and that blamed Aboriginal victims. These neutralisations fell away quickly when they sat in the circle with victims and elders. These are the senses in which restorative justice theory married to the economics of responsive regulatory theory can offer a better explanation than the narrowly economic theory of deferred prosecutions in Leone et al. (2021).

By the late 2000s, however, conversations within our ACCC reformers network were about enforceable undertakings becoming a soft option in the hands of many Australian regulators who allowed defendants to get away with saying, ‘We didn’t do it, but we won’t do it again.’ This became a national conversation when the Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry found in 2019 that enforceable undertakings negotiated by ASIC and the

Australian Prudential Regulation Authority had been consistently soft on bank criminality. Dukes et al. (2014) extended this critique to US corporate prosecutions concerning our longstanding work on corporate crime by Big Pharma. We found that recidivist pharmaceutical giants showed a pattern across four decades of paying bribes both inside the United States and globally and committing other serious corporate crimes, while settling with prosecutors one corporate integrity agreement, then another, then another. They scandalously breached its intent each time without being prosecuted.

Much as I continued to be attracted to restorative corporate justice, an inescapable conclusion was that the view had become widespread that corporations were ‘too big to fail’ or nail: law enforcement was captured by the concern that big banks must survive for the sake of the stability of the financial system and Big Pharma recidivists like Pfizer must survive because they hold patents to lifesaving drugs, as we saw with the huge contribution Pfizer made to building on government-funded university research on Covid-19. Pfizer needed to be a ‘fit and proper person’ to participate in government pharmaceutical benefits programs. Pfizer has long been by far the biggest, most important and most politically influential pharmaceutical corporation in the world. It negotiated one corporate integrity agreement with a restorative US state, offended again and negotiated a new corporate integrity agreement and then, when further offences were revealed by very senior Pfizer whistleblowing, US prosecutors negotiated a third corporate integrity agreement (Dukes et al. 2014: 339) and then a deferred prosecution agreement over alleged corrupt practices in eight countries (Paul Hastings 2012). In 2020, the US Department of Justice Foreign Corrupt Practices Unit opened yet another investigation over new possible breaches of this agreement in Russia, then China (FCPA Professor 2020). This followed another renewed line of litigation against Pfizer by US service and civilian personnel killed or wounded in Iraq and their families, alleging that corrupt payments by Pfizer to the Jaysh al-Mahdi terrorist group helped fund this group that attacked them (FCPA Professor 2020).

In the United States, there was a particular political history that drove this concern about enforcement capture. After Arthur Andersen was convicted and then had this conviction overturned by the Supreme Court for its role as the auditor of Enron and other corporations bankrupted during the stock market crash of 2001, Arthur Andersen itself effectively collapsed because of the adverse publicity surrounding the indictment, with 28,000

US employees losing their jobs. George W. Bush's administration drew the lesson that it was unfair that these innocent Arthur Andersen employees had lost their jobs. Australia had an identical debate when thousands of innocent Australian Arthur Andersen employees lost jobs against the background of some Arthur Andersen responsibility in major corporate crimes of 2001—in particular, the jailing of the CEO of the Australian insurance giant HIH.

Deferred corporate prosecutions were rare before the Arthur Andersen collapse. Coffee (2020: 38) found 419 deferred and non-prosecution agreements between 2002 and 2016, but only 18 in the previous 10 years up to the Arthur Andersen case. This section considers the possibility that deterrence may have been enhanced overall, at least in the terms that matter in the analysis of this chapter, and at least until the arrival of the Trump administration. That was because Garrett (2007: 855) found that prosecutors were laying charges in larger numbers of cases in the early twenty-first century than in the twentieth century; it is just that the increase was in deferred rather than completed prosecutions. Alexander and Cohen (2015) concluded empirically that the rise of deferred prosecution agreements has not suppressed other forms of corporate liability. Corporate cooperation with the individual accountability aspects of expanded deferred prosecutions has also meant that while there are reduced convictions of corporations, there is a small contribution to increasing convictions of individuals for corporate crimes in the United States (only 414 individual prosecutions across 306 deferred and non-prosecution agreements) (Garrett 2015: 1791) and some increased convictions in other countries with foreign corporations (Garrett 2014), plus some leveraging of foreign law reforms that could enable expanded future global enforcement (Garrett 2011: 1852). After the Global Financial Crisis, the Obama administration persisted with the shift from corporate crime prosecution to deferred prosecutions combined with compliance agreements. There were many unconscionable Obama administration failures to prosecute. Dukes et al. (2014: 185) wrote:

Restorative justice is important when it can lead to Corporate Integrity Agreements that genuinely confront and transform cultures of manipulation. Yet Corporate Integrity Agreements are not at present very searching;¹⁰ they fail to confront corporate cultures of manipulation. To date, they are no more than a tiny

10 This analysis cited the work of Ford (2008), Ford and Hess (2009, 2011) and Hess and Ford (2008).

step towards crafting a less manipulative industry that respects the spirit of the law in preference to gaming it. This is the reason [for taking] on the challenge of reforming Corporate Integrity Agreements by embedding them in a more robust framework, comprising tougher hybrid public and private law enforcement, restorative justice and transparent reporting and accountability for corporate integrity that transforms manipulation.

Changes to the board and the top management team are very common in deferred prosecution cases because of the adverse publicity in the financial press and as a consequence of the enforcement action, whether it is a prosecution or a deferred prosecution agreement. Corporate monitors were also appointed in 30 per cent of deferred prosecution cases (Arlen 2016). Regulators also often mandated corporate monitors pursuant to civil enforcement action. A chief compliance officer was often appointed to the top management group as part of the mandate. It may be that what Brandon Garrett (2007) called 'structural reform deferred prosecutions' that pursue 'deep governance reforms' indeed can transform. Garrett (2007: 855) illustrated transformation with the 2005 KPMG agreement to shut down its entire private tax practice, cooperate fully in the investigation of former employees and retain a former SEC chairman as an independent monitor for three years to oversee an elaborate corporate compliance program. The next chapter argues that the most important of such deferred prosecution undertakings go to corporate self-incapacitation rather than deterrence and that this is where their preventive potential mainly resides.

At least when it comes to big banks and Big Pharma, my professional experience of these organisations since the 1970s is that they have often remained recidivist criminalised corporations. If it is right that many of the organisations subject to deferred prosecutions that demand deep structural reforms are firms systemically criminalised across more than one kind of corporate crime, the governance and compliance reforms mandated by the deferred prosecution agreement may shut down other kinds of crime that have nothing to do with the offence for which the firm has been charged.¹¹

11 This was an early discovery of Stanley Sporkin's voluntary disclosure program after the Lockheed bribery scandals (Coffee 1977, 1981; Braithwaite 1984). External monitors of companies that had off-the-books slush funds for paying bribes found that criminal executives also used those accounts to rip off their company. For example, while Adnan Khashoggi may have persuaded ministers in foreign governments to buy Lockheed aircraft with bribes, he also gamed its slush funds to perpetrate massive fraud against Lockheed (Fisse and Braithwaite 1983, 1993).

If all this is so, why do CEOs not rationally defend this loss of sales and the sustained hit on the stock price by opting for a prosecution? One hypothesis is that CEOs do not defend the rational interests of the firm because cooperating with the prosecutor may be one path to their own survival. A CEO who keeps control during a negotiated settlement is likely to protect themselves and their board chair from losing their jobs and from prosecution by cooperating constructively and offering a lot of reform and repair to make the regulator, the prosecutor and civil society critics in the media happy enough. Even if a senior individual is going to be prosecuted, by cooperating to stay in control of events, the CEO can wield power to ensure the fall guys are vice-presidents responsible for going to jail (Braithwaite 1984; Garrett et al. 2019). Prosecutions are also expensive, and some ethical CEOs genuinely prefer to see money spent on victims rather than putting that money into the pockets of lawyers. CEOs also like negotiated settlements because they do not drag out as long as prosecutions, ending the distracting, debilitating uncertainty in their lives and in the market from high-profile enforcement. The deal can also do something to end the reputational damage quickly, which Fisse and Braithwaite (1983) found to be important to CEOs for its own sake, independently of the financial consequences of a reputational hit.

In the very first restorative justice case at the ACCC 30 years ago, the CEO refused to cooperate with the restorative justice process. The ACCC widened the restorative justice circle to include the chair of the board, who fired the CEO and then agreed to much more formidable undertakings than would have been imposed by a court (Parker 2004). Regulators grasped from the beginning that the rational CEO is a cooperator who, if pressed hard enough by victims, activist NGOs and the regulator, can give up a bigger financial loss in an enforceable undertaking than the maximum financial penalties in the law. Of course, that does not happen when the regulator is captured and victims and NGOs are quiescent, as predicted by responsive regulatory theory (Ayles and Braithwaite 1992: Ch. 3). There have been countless occasions during this century in the United States and Australia when this capture and quiescence have been present.

Alexander and Arlen (2018) concluded that convictions do not increase reputational damage beyond that imposed by deferred prosecution agreements, as long as the conviction does not reveal extra information about the firm's riskiness. This is true, among other reasons, because reputational damage does not depend on a plea of guilty. Extremely

large agreed payouts to harmed victims, a restorative justice of corporate contrition and revelations of facts to the media can do the reputational damage. These factors can also do reintegration through reform just as well, or better, than guilt conferred in the dock.

A final important datum from a restorative justice point of view is that corporations convicted by a judge in the United States pay an average of \$3 million in restitution to victims; in deferred prosecution agreements negotiated by prosecutors, the average is \$94 million, though this partly reflects the fact that larger corporations are more likely to receive deferred prosecution and smaller companies more likely to be convicted (Garrett 2014: Ch. 5). On average, victims get much more from civil suits arising from corporate crime than from public enforcement (Garrett 2014). All this says something shocking about how dismissive extant criminal sentencing is of victims' rights to restorative justice. Better integration of restorative victim compensation and prosecutorial punishment is another missed opportunity in how restorative justice could contribute to corporate criminal deterrence. This is not the most important defect of corporate criminal law from a restorative justice point of view. Much more attention to empowering victims with a voice, deep listening, apology, healing and prevention of further harm is required.

The largest criminal fine in Garrett's (2014) dataset was US\$1.26 billion against BP for the Deepwater Horizon oil spill. Probably more than \$28 billion was ordered in a combination of \$4.5 billion in civil penalties paid to the Justice Department and the SEC, civil suits or voluntary payments in compensation or for cleanup before this was demanded by any prosecutor or judge. Much of this is the old Waldman (1978) effect of delivering corporate self-enforcement to avert conviction or soften public enforcement. The US courts were stunningly kind and gentle to Halliburton and extremely tough on BP, if not as tough as BP was on itself. More importantly from a restorative justice point of view, while no great transformation in the corporate conscience of Halliburton has occurred, BP has sought to sell the story that it will transform itself from a carbon Goliath into a renewables David, committed to carbon neutrality by 2050, to become the oil major with the most transformative vision for the planet (Reed 2020). BP in turn sought to survive by filing \$40 billion in suits against the rig owner, Transocean, the rig cementer, Halliburton, and the blowout-preventer manufacturer, Cameron International. Pre-emptive self-punishment is also fundamental across all the data discussed. This involves investment in new compliance systems, appointing new

chief compliance officers and independent monitors of reform and firing senior managers. So much of this is done defensively as self-incapacitation (Chapter 10) in advance of demands that the self-punishment be done.

The most important point from an Australian restorative justice perspective about the Deepwater Horizon and Arthur Andersen cases is that cosmopolitan restorative justice in Australia could have prevented these catastrophes before they befell the United States. This is argued in the next chapter in relation to the power for the global corporate self-incapacitation of Halliburton in particular that was in the hands of Australian law enforcement. The Timor Sea oil spill, which was uncappable for 75 days and which occurred just a year before the 86-day Deepwater Horizon spill, which was uncappable for the same reasons at the hands of the same cement base contractor, Halliburton, should have produced a cosmopolitan restorative response for justice for future victims. Australian environmentalists should have demanded that the Australian regulator or courts require corporate monitoring reports of the cementing of all oil rigs around the world undertaken by Halliburton. Chapter 10 argues that the evidence is clear that this would have revealed a worldwide pattern of catastrophic risks with deepwater wells that were screaming to be fixed. Chapter 10 argues that, likewise, in the late 1990s, Australian regulators, particularly the Australian Taxation Office, were detecting a catastrophically criminal transformation of Arthur Andersen that might have catalysed cosmopolitan demands for global compliance and culture change, including at its Chicago headquarters.

One puzzle is why judges seem not to be as creative in their sentences as prosecutors working with regulators to impose compliance monitoring under deferred prosecution or non-prosecution agreements. There is no reason the judiciary could not become more creative custodians of corporate crime prevention in their sentencing.¹² One view is that law reform should force the judiciary into this role by requiring them to oversee the justice of both prosecution and non-prosecution agreements and the enforced self-regulation imposed by so many Australian regulatory agencies through enforceable undertakings. Then, perhaps, this limitation of corporate sentencing is because judges are not bureaucracies. Regulatory agencies (such as the Environmental Protection Agency with environmental crime and the Internal Revenue Service with tax crimes) working with prosecutors' officers have the better bureaucratic capacities

12 I am grateful to Brent Fisse for posing the question to me in this way.

to check the reports of corporate monitors, vet the suitability of the new chief compliance officer and oversee the rigour and transparency of their work in restoring integrity to the corporation. Moreover, there is the danger that activist judges who mimic regulatory bureaucracies will be accused of overreaching to usurp the policymaking responsibilities of elected officials, prosecutors who are accountable to elected officials but professionally independent, a politically accountable but professionally independent civil service and regulatory commissioners in the separation of powers (Baer 2016). Reasonably fearing this accusation, judges may always be too timid, too stretched, for the magnitude of a challenge so vast that they can never make the inroads required as the planet burns and financial systems unravel.

With deferred prosecution agreements and restorative corporate justice, the justice principles still seem the same as those Fisse and Braithwaite (1993) articulated. Most fundamentally, all who are responsible should be held responsible, be they individuals, firms or subunits of firms. That does not mean judges sentencing all of them to prison. If we have a president who gropes staff or a professor who gropes students, likely the university is responsible, the department is responsible, the head of the department is responsible and the groping professor is responsible. Fisse and Braithwaite (1993) have a lot of useful things to say about how to guard against scapegoating by powerful CEOs and how to hold CEOs individually accountable in appropriate ways while giving them credit for cooperation and for reactive acquittal of their fault (Fisse 1982). CEO fault is almost always something that societal collective efficacy demands as a remedy under Fisse and Braithwaite's accountability model—at least, managerial fault for the failings of operating procedures in a corporation they lead, but often criminal fault and civil liability under regulatory laws as well.

Given the power of CEOs to orchestrate smokescreens of diffused accountability to absolve themselves, societies do need to discuss the proposal of Harvard law professor Elizabeth Warren. She proposed during her 2020 presidential campaign that negligence should be the standard for CEO criminal liability inside criminalised banks that have catastrophic impacts on economies. Perhaps, however, Warren's objectives can be accomplished by tweaking what criminal recklessness means in the context of the power of a bank CEO. Perhaps given the pivotal importance of both CEOs and board chairs, remorse and accountability at those levels should also be encouraged by criminal law reforms that

require them normally to be in court for corporate criminal sentencing and at the press conference when the results of a deferred prosecution agreement are announced.

Corporate crime enforcement was wound back dramatically under President Trump. In the 15 years to 2017, aggregate corporate criminal fines increased from less than \$1 billion to \$10 billion a year, but they fell off a cliff in the next two years to one-fifth of what they were (to \$2 billion) (Garrett 2020: 116). Corporate convictions, convictions of individuals for corporate crimes and deferred prosecutions all declined during the Trump administration. Under President Biden, America could do worse than return to building on the limited progress that was being achieved through deferred prosecutions and *qui tam* actions for whistleblowers during the first decade of this century.

Misplaced optimism here could be a risk because the track records of the Clinton and Obama Democratic administrations were of beating the enforcement drum in press releases but muffling it during corporate negotiations that became politicised at times. The trajectory of the data on the stringency of deferred prosecution agreements post Enron suggests the same pattern of evolution towards a market in lemons described for Australian enforceable undertakings since 1990. Leone et al. (2021) created a cooperation score that combined corporate volunteering for self-investigation, timely reporting, prominent disclosure and the replacement of executives. Consistent with the Australian history, for the period 2002–10 (under the administration of George W. Bush), deferred prosecution negotiations left firms worse off: a 1 unit increase in a firm's cooperation score increased the probability of enforcement by 4.2 per cent and increased penalties by \$2.04 million. This result reversed during the Obama period of 2011–14, with 1 unit of higher cooperation reducing the odds of enforcement by 4.6 per cent and resulting in \$2.55 million less in fines (Leone et al. 2021). Firing allegedly culpable executives was particularly rewarding for firms during the Obama era, raising the scapegoating concerns that were such an issue in Fisse and Braithwaite (1993). Leone et al.'s (2021) results were consistent with an earlier study by Files (2012) that found cooperating with the SEC in deferred prosecution negotiations made firms worse off *before* 2010.

Untangling recursive causality in these data is difficult. Nevertheless, the data, combined with the Australian historical experience of enforceable undertakings, are sufficient to conclude that the pursuit of some static,

evidence-based set of optimal deterrent policy settings is folly. Much depends on how tough regulators and prosecutors are in deferred prosecution negotiations. That constantly shifts with the political winds, the excellence of regulatory leadership and shifts in corporate cultures of responsibility. The interactions among corporate, CEO and lower-level sanctions that may target executive scapegoats are complex, as are interactions between deterrence and preventive incapacitation. This is no warrant for nihilism. It justifies carefully monitored dynamic responsiveness, more genuinely restorative and responsive participation of third parties (particularly victim representatives) in the sanction negotiations, rather than state–corporate deals behind closed doors. The imperative remains unrelenting citizen and social movement vigilance against regulatory capture and corruption (Ayres and Braithwaite 1992: Ch. 3).

It also justifies consideration of further strengthening enforcement with Coffee's (2020) ideas on equity fines and the privatisation of corporate criminal enforcement, but informed by a more restorative and responsive philosophy of prevention and punishment. Coffee is surely right that access to potent equity fines would strengthen the negotiating clout of the public interest against corporate power in deferred prosecution negotiation. Baer (2016) may be right that the challenge of prosecuting every guilty corporate criminal to fix the crisis of 'too big to fail, too big to jail' is a mission 'too vast to prevail'. Yes, we can spend a good bit more on corporate prosecutors, but funding more street-level regulatory inspectors, fraud examiners, environmental NGOs and activists in networks like Citizens for Tax Justice may be a higher priority for reducing the suffering and domination caused by corporate crime than funding more lawyers and building more prisons.¹³ More conceptually, my hypothesis is that the collective efficacy, the bridging and linking social capital of a society against corporate crime, may be what matters most.

13 As Coffee (2020: 93) himself concedes: 'Civil enforcement dwarfs criminal enforcement, whether in terms of manpower allocated, aggregate damages collected, or numbers of actions brought.' With banks, Coffee (2020: 93) points out that the six worst bailed-out US banks had experienced 'more than 350 "major legal actions" that had imposed approximately \$182 billion in sanctions and settlements'. While the Global Financial Crisis did not produce major successes from criminal prosecutions until the final years of the Obama presidency, in the years immediately after the crisis, civil regulatory actions against banks peaked at 2,208 per year (Coffee 2020: 94). In turn, the financial recoveries from civil and criminal regulatory penalties combined were dwarfed by recoveries from securities class actions (Coffee 2020: 104). In my terms, regulators, internal compliance actors and civil society actors are the ones who deliver most of the detection and the diversity of control mechanisms in Table 6.1.

Deterring police departments as criminal organisations

Reflexivity is in scarce supply in criminology. Obviously, we must study police as both preventers and perpetrators of crime. Chapter 8 discussed the implications of Lawrence Sherman's (1978) evocative study of police scandal and reform. One reason prosecutors are reluctant to prosecute police is that they are hard targets who are expensive to nail. They are sophisticated witnesses, adept at orchestrating reasonable doubt and smokescreens of diffused accountability for wrongdoing. Police chiefs who seem corrupt to many in the community rarely go to jail; police constables who seem to many in the community to have murdered suspects are rarely convicted of murder. It makes little sense for the community to even attempt to punish itself by fining police departments that rely on the community's taxes. Police departments may be too big and too politically connected to fail, and too street smart to nail.

Nevertheless, police organisations are like other public sector organisations such as universities in that their leaders care about their reputation for its own sake, quite independently of any financial implications of a reputational hit. Chappell (2017) found that consent decrees settled with 23 agencies subject to US Department of Justice litigation for police misconduct produced a 23–36 per cent reduction in subsequent filings for further alleged civil rights violations. It is impossible to say whether this might be a deterrence effect mediated by police leaders' concern for their reputations. Based on Sherman's (1978) research, we might conjecture that reform was more likely to be a self-incapacitation effect negotiated under the consent decree. Internal police integrity testing was particularly important in the corruption scandal and reform cycles studied by Sherman. These are questions that are hard to answer, but we might do best to consider criminal police organisations as just one kind of criminal organisation. We should be open to the frame that criminalised police are just one bit of criminalised states, as we consider the potential of enforced organisational self-incapacitation in the next chapter.

Conclusion

Inexorability is a core principle of minimally sufficient deterrence: pursue inexorable consistency of detection and disapproval of predatory crime. This implies fusing the debate on dynamic concentration of deterrence with the debates about less prison and more and better street-level state monitoring and collective efficacy in civil society.

The move away from the nihilism about policing effectiveness prevalent at the time of the Kansas City Preventive Patrol Experiment has lessons for criminologists (Nagin et al. 2015); as has the shift from nihilism about rehabilitation prevalent at the time of the 'Nothing Works' slogan (Lipton et al. 1975). Policing and rehabilitation are useless or dangerous only if they are unresponsively deployed. For example, evidence-based refinement of the responsivity of rehabilitation can improve the menu of options in the pyramid of supports in Figure 9.1 (Andrews and Bonta 1998, 2010; Manchak and Cullen 2015). Weisburd et al.'s (2017) systematic review of 118 separate systematic reviews finds that a wide variety of interventions are quite effective in reducing crime when they strengthen what have here been called human capital, social capital, recovery capital and restorative capital, and a wide variety of other interventions to improve policing, diversion and mentoring or to close off criminal opportunities. The exception to this widely variegated pattern of greater and lesser effectiveness was the ineffectiveness or harmfulness of punitive sentences. Developmentalists convincingly showed that social support is important to crime prevention long before the first offence occurs (Cullen 1994). This is a vital piece for any integrated theory of crime prevention. Deterrence is therefore far from the most important element of a sophisticated strategy to protect citizens from crime and guarantee their freedom. Deterrence is less important than sound management of anomie and building plural forms of social capital, as discussed in previous chapters. Deterrence is less important than incapacitation, especially of the crimes of the powerful, as discussed in the next chapter.

'Less prison, more police' (Durlauf and Nagin 2011a) is not convincing as a slogan; nor is 'defund the police'. Having more police is an unpersuasive idea when so much policing in the United States and Australia is racist in ways that reduce freedom and increase crime. Meta-analyses such as that of Pratt and Cullen (2005) show that increased funding of police, police per capita and arrest ratios are at the bottom of their list of macrolevel

predictors of crime, often engendering defiant backlash that makes crime worse (Sherman 1993). Having more police is only a good idea when policing is not racist, is evidence-based and steeply reduces arrest as its default strategy. In the end, this book also argues that transformed police, and more of them, could be a small part of a shape-shifting reform for reducing the environmental destructiveness of economies that avert the crises that most threaten freedom and violence (Chapter 12). Evidence-based policing can be part of a strategy for achieving economic growth in human services rather than in consumer durables. More police can also save surprisingly large numbers of lives when deployed to UN peacekeeping operations (Hultman et al. 2013).

A better slogan than ‘Defund the police’ on my analysis is ‘Less prison, less arrest’. One reaction to many of the police killings that motivated the ‘defund the police’ movement was: ‘Why were the police seeking to arrest this citizen in the first place, why did they pull out a gun when there was some resistance and why were they armed in the first place?’ Bad criminological ideas like broken-windows policing that have been implemented in racist ways have contributed to the overuse of arrest to the point where ‘in our society liberty is not the norm and detention prior to trial or without trial is not the carefully limited exception’ (VanNostrand and Keebler 2007: 23). Police arrest policies have paved this path to tyranny.

Braithwaite’s (1989) theory argues that when police are reintegrative, they can reduce crime; when they are stigmatising and violent, they increase it, which is why having more police does not currently lead to less crime. Kennedy (2017) rightly argues that there are US police departments that have reduced the number of arrests, reduced the number of complaints against them, reduced incarceration rates and reduced crime through evidence-based policing. They are exceptions at this point in history. Minimally sufficient arrests are a path to crime reduction and to enhancing freedom. Engel et al. (2017), for example, show persuasively how the Cincinnati police accomplished this after the Queensgate Correctional Facility was closed. Street-level police were persuaded to a cultural transformation whereby arrest should be used sparingly. Violent crime, arrests and imprisonment were simultaneously reduced by a combination of the evidence-based policing strategies discussed herein: hotspot policing, dynamic concentration of deterrence of the Operation Ceasefire variety, problem-oriented policing and expanded welfare resourcing of partnerships with social service and health agencies.

Suggestive evidence has been introduced that an inexorably supportive firm hand might help in preventing crime, in preventing the collapse of welfare states that struggle to deter corporate tax evasion and in addressing many other challenges of crime control. The white-collar crime piece of this is important. Any theory of crime that provides an account of crimes of the powerless but not crimes of the powerful is troubling and, indeed, misleading. It might be credible as a theory of something more specific than crime. Moreover, the dominance of theories in criminology that fail this test means criminology buttresses oppression when it normalises prisons that hold tiny numbers of wealthy white criminals.

The evidence adduced in support of minimally sufficient arrest, minimally sufficient prosecution, minimally sufficient imprisonment and minimally sufficient deterrence is no more than suggestive. It is common for criminological theories to have something going for them while being wrong in most contexts. Until minimally sufficient deterrence is subjected to an array of different kinds of empirical investigations, this may be as true of it as it is of the theories of passive deterrence that currently dominate thinking. I have attempted to show that minimally sufficient deterrence has promise as a strategy for moving from passive to dynamic deterrence because it starts from what we already know about deterrence and defiance and because it integrates insights from other relational theories that each enjoy a body of empirical support. These are theories of social support (Cullen 1994), social capital (as discussed in Chapter 7), responsivity (Andrews and Bonta 1998, 2010), responsive regulation (Braithwaite 2021f), sharpening the Sword of Damocles (Dunford 1990; Sherman 1992, 2011), dynamic concentration of deterrence (Kleiman 2009), shame and pride management (Ahmed et al. 2001), combined with indirect reciprocity (Berger 2011; Nowak 2012), and motivational interviewing (Lundahl et al. 2010). The imperative, grounded in complexity theory, for abandoning applied social science that tests specific parsimonious theories in favour of applying meta-theories has been explored. These are theories about how to organise multiple theories and meta-strategies—strategies about how to sequence many strategies.

While minimally sufficient deterrence is based on what we know about deterrence and defiance, that knowledge base has wide gaps of complex unknowns (Braithwaite and D'Costa 2018: Ch. 12). The future gap-filling research agenda can be framed under the seven policy principles of minimally sufficient deterrence:

1. Escalate enforcement: Display intent to progressively escalate a responsive enforcement pyramid that involves progressive escalation of sanctions for wrongdoing and support for social responsibility.

This has been the heartland research priority of Valerie Braithwaite's and my research group since 1980—for example, see Braithwaite's (2021f) review essay and more than 100 empirical evaluations of the application of responsive regulation to tax compliance by the Centre for Tax System Integrity (ctsi.org.au/; more broadly, see johnbraithwaite.com/responsive-regulation/).

2. Inexorability: Pursue inexorable consistency of detection of predatory crime. Communicate inexorable community commitment to stick with social support for those struggling with problems of lawbreaking until the problems are fixed.

Critical research contributions here bring together the established agenda of measuring the effects of perceived certainty of detection with the belief that supporters of offenders will deliver them unconditional support, sticking with offenders' problems until they are fixed. While increasing consistency of detection will increase deterrence, police being everywhere at all times risks undermining legitimacy and motivating defiance, especially when some police are stigmatising or inflame racial injustice. Lawrence Sherman has coined the idea of a sweet spot of intensity of just enough deterrence through police presence at hotspots. Gibson et al. (2017) found such an optimal sweet spot of minimally sufficient patrols in Merseyside, in the United Kingdom. Though it is well established that intensive patrolling at hotspots can reduce crime (Braga et al. 2014), Gibson and her colleagues are the first to explore the possibility of reducing the intensity of hotspot patrolling without increasing crime, perhaps even reducing it somewhat through optimising each sweet spot. This work opens a path to understanding cost-effective, minimally sufficient patrolling.

3. Escalate social support: With repeated offending, increase social support. Even when there is escalation to a last resort of severe incapacitation, escalate social support further. Keep escalating social support until desistance is consolidated.

Perhaps the most critical research needed here is macrosociological and economic work on strategies for sustaining a more credible welfare state, a topic re-joined in the final chapter. It is feasible to be politically effective in struggling for a return to progressively improving the welfare state.

4. Sharpen the Sword of Damocles: Cultivate the perception that ‘trouble hangs inexorably over my head; they want to support me to avert it’.

Here, the ‘less prison’ research agenda shows the kind of work that illuminates Sword-of-Damocles possibilities (Sherman 2011). This is illustrated through Slothower et al.’s (2017) ‘West Midlands Police experiment, Offender Management by Turning Point (Deferred Prosecution with a Plan)’. Random assignment to deferred prosecution combined with social support substantially reduced criminal harm (by 34 per cent) though not the incidence of crime, reduced the cost of the justice system and increased victim satisfaction with outcomes when compared with prosecuted cases. Moreover, the deferred prosecution ‘did something’—something constructive that reduced costs, averting a world in which ‘nothing’ happens until one day a lot happens. This lot that happens then seems arbitrarily harsh. From the perspective of this chapter, a tempered ‘something’ that happens is also a better approach to constructive structural sharpening of the Sword of Damocles.

5. Dynamic concentration of deterrence: Focus deterrence on a line that should never be crossed after an announcement date. Then progressively lift that line in high-crime contexts, raising our expectations of socially responsible citizens.

Research in this tradition led by David Kennedy and Mark Kleiman has not been linked to evidence-based learning on restorative justice and responsive business regulation, nor to the dynamic concentration experience of international peacekeepers regulating war zones and negotiating gang surrenders to create peace zones. A more interdisciplinary research imagination is required to see the complex of strategies, including escalated social support and reconciliation, to embed dynamic concentration. This can increase the effectiveness of deterrence and justice. Future research must distinguish static, focused deterrence effects from dynamic concentration effects.

6. Community engagement: Engage the community with offenders in widening restorative conversations that educate about the shameful of criminal predation for the many who participate in the conversations. Avert stigmatisation.

The research required here includes the intersection of work on community engagement with crime control (for example, Sampson et al. 1997; Pratt and Cullen 2005; Odgers et al. 2009) and on the Connectedness, Hope, Identity, Meaning and Empowerment (CHIME) conclusion reached by Leamy et al. (2011) in their review of recovery capital research (Best 2017). The CHIME conclusion is that connectedness, hope, identity, meaning and empowerment are needed for freedom as capability (Sen 1999), for recovery from problems such as drug addiction, alcoholism, suicide attempts and arrest. It is important to integrate the best psychological and criminological research on pride and shame dynamics and on shame acknowledgement as offenders re-narrate their lives (Leach and Cidam 2015; Spruit et al. 2016; Braithwaite 2020c). The community best learns the shameful of corporate crime through media coverage of stories of corporate harm and restorative contrition, apology and repair.

7. Modesty: Settle for the modest general deterrence delivered by this shameful and a minimal number of cases that escalate towards the peak of the enforcement pyramid.

This is the 'decremental' research strategy commended by Braithwaite and Pettit (1990) for republican freedom and criminal justice. It means evaluating research on how low imprisonment can go without crime beginning to increase. When we have no choice but to lock up extremely dangerous people, we can be justifiably pessimistic that this will deter those specific people when they are released. Yet others noticing that imprisonment does sometimes happen may deliver a modest quantum of general deterrence of the rest of the population. Braithwaite and Pettit's (1990) decremental research agenda has gone nowhere in 31 years. No country pursued progressive reductions of imprisonment rates until evidence emerged that serious crime problems were the result. This is a measure of how wide the gap is in every country between minimally sufficient deterrence and criminal justice policy.

This text is taken from *Macrocriminology and Freedom*, by John Braithwaite, published 2022, The Australian National University, Canberra, Australia.

doi.org/10.22459/MF.2021.09