Simon Bastow

Governance, performance, and chronic capacity stress in public policy systems

Working paper

Original citation:

Originally available from LSE Public Policy Group

This version available at: http://eprints.lse.ac.uk/53749/
Available in LSE Research Online: October 2013

© 2013 The Author

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.
GOVERNANCE, PERFORMANCE, AND CHRONIC CAPACITY STRESS IN PUBLIC POLICY SYSTEMS

Simon Bastow

LSE Public Policy Group
London School of Economics and Political Science

Contact s.j.bastow@lse.ac.uk

Public policy systems often betray signs of capacity stress. Supply factors seem to lag behind demand, and as a result these systems sustain what appears to be chronically sub-optimal performance neither completely excelling nor completely failing in what they do. How should we interpret these syndromes? Although competing theoretical explanations can shed partial light, potential ‘blind spots’ between them mean that they are not able to capture capacity stress in its entirety. In this paper I introduce the concept of ‘chronic capacity stress’ (CCS) as a holistic governance-style approach. I show how capacity stress is closely associated with four interrelated factors: 1) misalignment between policy goals and the external environment; 2) cultures of coping and adaptation; 3) direct and benign resistance to change; and 4) sustained obsolescence and inefficiency. The CCS heuristic provides a useful way of understanding dialectical dynamics of stress and equilibrium in UK policy systems.
Frequently the disruption of a political system is not that complete; the stress is present even though the system continues to persist in some form.

David Easton

For many good reasons political scientists have devoted attention in recent decades to understanding the extremes of public sector performance. At the crisis end of the spectrum, concepts such as ‘policy disasters’ (Gray & T’hart, 1998), ‘policy fiascos’ (Bovens & T’hart, 2005), ‘service delivery disasters’ (Dunleavy et al., 2010), or ‘government blunders’ (King & Crewe, 2013) have offered useful insights into why things have a tendency to go badly wrong. At the success end, there has been less systematic analysis (McConnell, 2010), but interest none the less in the application of aspirational management concepts to the public sector (Osbourne and Gaebler, 1992; Talbot, 2010). Although focusing on these extremes can be helpful in understanding how policy systems work (or do not), it should not distract from the fact that many show signs of operating in a qualitatively complex middle ground between the two extremes of complete success and abject failure.

If we look carefully and long enough at large policy systems, particularly those that fulfil core functions of state, they often betray signs of continual capacity stress. The perception is that the demands and expectations that are put on these systems by governments and society are greater than their capacity to respond. Supply of outputs (and outcomes) appear to lag persistently behind perceived demands whatever they may be. Symptoms include continual backlogs, insufficient staffing and resourcing, crowding and congestion, or not having the right equipment in the right place at the right time. Stress may be insufficient to bring about acute collapse of the system (although from time to time acute problems may flare up). But it is significant enough to stop the system from performing to its optimal potential. In short, these systems find ways of absorbing stress, adapting or coping with it, and often normalizing it so that, as the quote above suggests, they persist in some form or other.

Competing theories of social behaviour however take us only so far in understanding this qualitatively complex middle ground. Reliance on particular
theories shed some light, but also perpetuate ‘blind spots’ in our overall understanding of how and why policy systems sustain capacity stress. Indeed, critics have expressed a need for more integrative and ‘negotiated’ theoretical approaches to understanding governance problems (Peters, 2010; Cairney, 2013a). The aim of this paper is to make a contribution along these lines by introducing the concept of ‘chronic capacity stress’ (CCS) that can help us to bring these potential blind spots into more direct focus. I define CCS as follows:

- A significant stress on the capacity of a policy system to respond to expectations and demands made on it;
- one that is sustainable over time (often long periods), and can be managed as such, and therefore becomes chronic;
- one that is a function of the way in which the system itself is governed and operates; and
- one that is simultaneously both cause and symptom of dysfunction in the system.

The concept itself has grown out of my research on overcrowding in the England and Wales prison system (Bastow, 2013a). This has been an archetypal problem of capacity stress for many decades, and it provided the detailed empirical basis from which I have developed the outlines of CCS (Bastow, 2012b). As I show here, this concept has more general applicability across different policy systems.

The paper is in three parts. First, I discuss three potential ‘blind spots’ that are sustained by axiomatic theoretical approaches. I explain how the problem of capacity stress draws attention to these blind spots, and why we need something like the CCS heuristic to neutralize their effects and to formulate tools for diagnosing stress-related syndromes in large policy systems. In the second part I discuss the four aspects of CCS and inherently compensating and countervailing relationships between them. These four component parts are as follows, 1) misalignment between goals of the system and external environment; 2) cultures of coping and adaptation; 3) benign resistance to change; and 4) sustained obsolescence and inefficiency. I draw on examples from the UK public sector, including immigration and border control, welfare administration, defence, and prisons. Finally, I illustrate how these
four aspects interact in compensating and dialectical ways, and the implications for these theoretical blind spots.

Three potential blind spots

Just as the brain renders invisible the physiological blind spot and gives the illusion that the visual field is continuous and complete, so the mythology of science has the function of hiding from view the holes in the fields of consciousness and rationality.

*William Byers*

In social science, dichotomies can often help researchers define the realm of all potential observable outcomes. Dichotomous pairs mark out this realm and, as the quotes suggests, give substance to the ‘illusion’ that the visual field (i.e. all observable outcomes) is ‘continuous and complete’. But they can also perpetuate blind spots. We may assume that pairs exist at either end of linear or continuous spectrums, on which it is possible to identify ‘blended’ points between both. The problem here is that complex realities often involve distinct qualitative characteristics of both at the same time, and not just a single average point between the two. It is necessary to make room for the fact that policy systems can display aspects of success and failure simultaneously, thus casting doubt over the assumption that it is continuous blended terrain from one extreme to the other. In this section, I discuss three potential blind spots.

The first blind spot lies in the relationship between economistic explanations of chronic stress through *over-supply* and *under-supply* of outputs. From the 1970s onwards, many critics diagnosed the syndrome of the UK state as one of chronic over-supply. This reflected a perception of largely unchecked and excessive growth in the size of the state, and a sense that it had become unmanageably large, top-heavy, and in need of radical downsizing (King, 1975; Rose, 1979). Public choice theorists of the era, notably William Niskanen (1971), provided theoretical
explanation for this perceived top-heaviness through the self-interested 'budget-maximizing' behaviour of bureaucrats (and, we assume, politicians). Such behaviour was key to understanding chronic over-supply of outputs by the state and public sector. Reining in the ‘maximizing’ behaviour of the state and its constituent actors was seen as constituent of a more sustainable balance.

As ‘new right’ ideologies took hold in the UK throughout the 1980s and beyond, waves of managerialist and market reforms swept through the state, largely in response to these perceptions of overweight regimes. Whereas critical focus had previously been on the tendency for self-interested elites to sustain over-supply, political scientists began to see the problem implicitly as one of chronic under-supply. Here these (same) bureaucrats could find ways of successfully extracting rent or ‘shaping’ their budgets so as to extract value from the system for their own benefits, and subsequently pass the costs of this extraction down to actors below them in the system (Tullock, 1967; Dunleavy, 1991). Elites had incentive to shape their own jurisdictions and budgets, extract value, and consequently squeeze the system below them in order to compensate for their own value-extracting activity. The analogy here is one of chronic tightening, increasingly forcing actors below to run at ever more ‘close-to-tolerance’ levels. This cumulative squeezing through the system sustained capacity stress through chronic under-supply.

This dichotomy, grossly simplified though it is here, raises questions for how we understand capacity stress. In economistic terms, we can envisage both concepts as lying on a continuum around a point at which supply and demand curves are in some kind of equilibrium. Either we are over-resourcing a prison system and need to find ways of trimming excess fat; or we are under-resourcing it and need to find new money and capacity. Logically there is one point somewhere in the middle at which the two intersect. The potential problem with this however is that such linear understanding distracts from the fact that it may be possible for policy systems to sustain qualitatively distinct elements of over-supply and under-supply at the same time. In systems under stress, we find signs of apparent under-supply as they are squeezed (particularly in an austerity era as resources are cut). Yet, paradoxically, the very fact that these systems are under stress means that they sustain inefficiency,
obsolescence, and redundancy, by being unable to innovate effectively or introduce long-term modernization. In short, they can sustain relative over-supply as well.

The second blind spot relates to a well-established axiomatic distinction between what Mary Douglas (1986) has called the ‘double-stranded view of social behaviour’ (p18) - transactional and cognitive explanations. The transactional strand relates to ‘individual utility-maximizing activity described in a cost-benefit calculus’, while cognitive strand relates to ‘the individual demand for order and coherence and control of uncertainty’ (p18), largely through development of institutions and constituent cultures. Although significant contributions have been made over the years to reconcile these strands (not least Scharpf (1997) from the cognitive camp or Besley (2006) from the economistic transactional camp), critics still point to the inherent difficulties of reconciling both extremes in ways that help to understand governance of complex policy systems. Peters (2010), for example, reflects that ‘governance approaches do not have any explicit mechanisms of integrating individuals and structures’ (p17) and suggests that ‘many approaches to social theory tend to focus on a single explanation or actor, rather than on how the possible explanations can be brought together in a more comprehensive explanation’ (p8).

Concepts over the years have helped to bridge this gap and explain how failure can become chronic in organizations. Meyer and Zucker’s (1989) ‘permanently failing organizations’ has much to offer in this sense, grounding its explanation of ‘failing’ in the strategic self-interest of influential ‘dependent’ actors, perhaps strong unions or professional groups, to perpetuate sub-optimal status-quo situations. The appeal of this concept is that it provides a specific mechanism through which organizations chronically under-perform. At the same time, the very concept of ‘permanent failure’ implies strongly fatalistic institutional culture (Hood, 1998), and this can have an important feedback effect on the way in which actors in the system view their predicaments and possibilities. If a system is labelled as ‘failing’, this must to some extent shape the content and limitations of strategic self-interest. There may be little point, for example, in senior ministers setting out to ‘reform the un-reformable’ when they can opt to ‘play a straight bat’ until they move up to another portfolio. Equally, there may be strong incentives for coping managers or staff to portray their bit of the system as being on the brink of collapse or crisis,
with a view to securing extra attention or resources. Hargrove and Glidewell’s (1990) ‘impossible jobs’ also depicts similarly deep-seated interaction between transactional and cognitive dynamics. In characterising the challenge for the top officials or commissioners in those systems, they suggest that ‘commissioners cope with an impossible job; they never master it or control it’ (p45). When a system is out of touch with exogenous pressures from its wider environment, a dominant response for actors responsible is to find ways of adapting and making things work.

The third potential blind spot relates to the dichotomy between deterministic and holistic explanations (i.e. parts versus whole). Political scientists seek to formulate specific and ‘testable’ hypotheses about causal factors behind capacity stress, and aspire to what King et al. (1994) describe as ‘parsimony’ of explanation. Yet, capacity stress, indeed any kind of stress, is often the outcome of compensating and dialectical dynamics. In order to see these dynamics working it seems reasonable that we should set our lens wide enough to see the whole system working (not just bits of it), and the more we venture towards holistic appreciation of the problem, the more we inevitably sacrifice parsimony for a ‘more general notion of maximizing leverage’ (King et al., 1994, p104).iiii The dichotomy between determinism and holism therefore requires that we have flexibility in the focal length of our evaluation, as well as specificity about the parameters of the system and actors that we are including in the frame.

Holistic approaches to understanding the characteristics of policy systems have not been lost on political scientists over the decades. Throughout the 1960s, systems-style explanations of political and social phenomena developed prominence in Europe and in the US. Crozier (1964), for example, identifies specific mechanisms inherent in bureaucracies that perpetuate inward-looking cultures, and lags in taking up what we acknowledge as standard modern practices in society. Crises, for Crozier, play a vital function in forcing these inward-looking systems to adapt. David Easton’s (1967) work on political systems examined how and why political systems are able to sustain and adapt to stress, both internally generated within the system and externally from its environment. As he put it:
Political systems accumulate large repertoires of mechanisms through which they may seek to cope with their environments. Through these they may regulate their own behaviour, transform their internal structure, and even go so far as to remodel their fundamental goals. Few systems, other than social systems, have this potentiality’ (p18)

Holistic approaches to public problems have come back into fashion in recent decades. Cybernetics (Dunsire, 1990), complexity theory (Teisman & Klijn, 2008; Teisman et al., 2009), and evolutionary theory (Cairney, 2013) have all been applied in explorative ways to public policy and governance. In different ways, these have drawn on a three-way relationship that is central to the holistic viewpoint:

A. prevailing values and goals of the system;
B. the way in which the system is set up to deliver on these objectives; and
C. the way in which these two factors are aligned with the system’s external environment.

Adapted in a strategic management context, Roberts (2004) explains the generic holistic challenge as ‘one of selecting the (long run) value-maximizing strategy for the particular environment and then creating the organization that will best realise it’ (p22). The central point for Roberts here is the quality of the dynamic equilibrium and complementarity between these three concepts. He argues that it is impossible for the system to maintain continually optimal alignment. The system as a whole will inevitably incorporate certain aggravating misalignments with its environment that put pressure or stress on its ability to perform. For him, and other management scientists, the best that can be hoped for is a kind of acceptably sub-optimal level of alignment over time.

It is therefore important to keep the analytical lens as wide as possible. After all, understanding how systems sustain stress requires that we are able to conceptualise the system itself and how it interacts with its environment. Stress is an artefact of the system as a whole, not simply bits of it. Sceptics may immediately criticize the ‘everything-and-nothing’ nature of this term (Phillips, 1976; Rosenberg, 1995), and question legitimately how we know which ‘bits’ to include and which not
to include. It is all the more important to be cautious in defining the parameters of the system, and to look for specific (and even testable) concepts or variables that explain how and why systems sustain capacity stress. In this sense, we are looking to be as deterministic as we possibly can be, while maintaining sufficient looseness in the overall approach that we are to feed actual empirical observation into the development of the overall model. To the development of the CCS model is where I turn in the next section.

Four inter-related aspects of CCS
In setting out these potential blind spots, it is not a case of accepting or rejecting these constituent theories. The dichotomies help us to envisage the parameters of the theoretical universe out there, and they must in this sense form the component parts for any theoretical understanding. But it is likely also that they only provide strands of insight into how the actual universe functions in reality. By introducing the concept of CCS, the aim is to push forward our understanding of how many large and complex policy systems ‘actually operate’, as Amann (2013) puts it, and provide a schema or heuristic for straightening out the distracting or distorting effects of these blind spots.

The development of the CCS heuristic is methodologically deductive and inductive. Indeed, this is a further blind-spot-inducing dichotomy that I have not yet mentioned, but it shows exactly how the constructed norms of methodological theory can often bear little or no resemblance to the way in which theoretical and empirical knowledge interact in real-life research projects. The theory discussed in the previous section provides the foundation for the CCS concept. In this sense, it builds on nothing else but existing theoretical knowledge. At the same time, however, its development is absolutely contingent upon in-depth empirical research into whole policy systems, and the aggravating and compensating dynamics that sustain ‘optimal sub-optimal’ equilibriums in them. We can only learn about this however by talking to actors in the system, understanding their predicaments, the extent of their constrained autonomy, and the overall outcomes that cumulatively emerge.
The root foundations for CCS are none the less deductive. In a sense, the order in which the three blind spots were discussed in the previous section depicted a ‘layering-up’ of theoretical explanations from individualistic to whole-system approaches. In developing the CCS schema, we can work deductively in reverse, setting out the holistic picture and digging down into the transactional and cognitive dynamics within. We pick up therefore where we left off with the three holistic aspects – goals, organization, and environment. I generate four relationships between them that together form the basis of a holistic appreciation of how a policy system operates dynamically in its environment. The schema is set out below in Figure 1, and includes these four alignments (or potential misalignments) that could conceivably aggravate stress in a policy system. These are indicated by the arrows and labels (a) to (d). I work through each of these four potential misalignments drawing on some brief illustrations from a range of UK policy sectors.

**[FIGURE 1 ABOUT HERE]**

**a) Value-based goals are misaligned with changes in the external environment**

When societies change, or at least the problems they present, we may expect the values and goals of policy systems to change with them. But this is not necessarily the case. Societal change, although it may evolve gradually over time, can often take policy makers and governments by surprise. It can happen imperceptibly and may only be brought to awareness of policy makers through periods of punctuated equilibrium, perhaps as ‘policy streams’ converge (Kingdon, 2003) or as external shocks trigger *ex-post* policy responses (Depoorter, 2006). At key junctures, governments and policy makers make sweeping changes to redress misalignment that has been allowed to build up. Alternatively, actors may be well aware of gradual change, but these may present long-term collective action problems and policy makers are able continually to avoid the issue and leave it for their successors to negotiate.

Value-based goals may also be too ambiguous or not robust enough to maintain sufficient clarity about what policies seek to achieve. Rainey and Chun (2005), for example, see ‘goal ambiguity’ as having a marginal undermining effect
on performance in large public sector organizations. But as Roberts and colleagues suggest, firms can often be high-performing simply by maintaining ‘optimal sub-optimality’ across competing goals, and the same presumably applies to public sector systems. The prison system, for example, has for many decades sought to find ways of reconciling self-stated goals around deterrence and reform (McConville, 1995a, 1995b). Similarly, the armed forces procurement logistics continually have to reconcile the need to have modern and functioning equipment ready on time for armed forces to deploy, while keeping procurement and inventory costs as low as possible.

Stress is aggravated if sudden changes in the external environment conflict with prevailing goals. For example, the unprecedented increase in the size of the short-sentence prison population in England and Wales prisons from 1993 onwards can be traced to more punitive sentencing by the courts (Hough et al., 2003). In 1993, the average number of prisoners sentenced to less than six months custody entering the prison system was around 22,000, yet by the start of the 2000s, this number had more than doubled to over 50,000. While the prison system was filling with relatively low-risk short-term offenders in need of targeted rehabilitative interventions, the political priorities of the system were focused on how to maintain security and ensure zero escapes (Cavadino and Dignan, 2003 and 2006). It was not until the mid-2000s that the policy elites in the prison system began to shape policy goals more explicitly towards this increasing short-sentence section of the prison population.

We see a similar lag in ability of governments and UK armed forces to move away from Cold War paradigms of conflict and procurement towards new forms of strategic engagement associated with a post 9-11 terrorist and 'rogue state' era. The following text is taken from the 2010 Strategic Defence Review:

Our Armed Forces – admired across the world – have been overstretched, deployed too often without appropriate planning, with the wrong equipment, in the wrong numbers and without a clear strategy. In the past, unfunded spending pledges created a fundamental mismatch between aspiration and resources. And there was a failure to face up to the new security realities of the post-Cold
War world. The Royal Navy was locked into a cycle of ever smaller numbers of ever more expensive ships. We have an Army with scores of tanks in Germany but forced to face the deadly threat of improvised explosive devices in Iraq and Afghanistan in Land Rovers designed for Northern Ireland. And the Royal Air Force has been hampered in its efforts to support our forces overseas because of an ageing and unreliable strategic airlift fleet. This is the result of the failure to take the bold decisions needed to adjust our defence plans to face the realities of our ever-changing world. (MOD, 2010, p4)

b) Organisational design of the system is not set up to deliver value-based goals

Whatever the goals of a policy system, there is pressure on the actors at all levels to deliver against those expectations. When the system is not sufficiently set up to meet those goals (either in perception or in actuality), then it is likely, as Hargrove and Glidewell remind us, that cultures of coping will persist. There is potentially complex interplay between strategic self-interest and cognitive cultural implications of coping. Actors in the system may see their own predicament as one of having to cope with the pressures imposed from above or externally, and there is often a permanent sense of fatalism that pervades the organization. Particularly in the types of agencies more commonly associated with ‘impossible jobs’, staff see their levels of professional autonomy as more constrained (CSPS, 2013). Yet officials are also keenly aware of the strategic benefits of perpetuating a perception that the system is (just about) coping on the brink of acute crisis. Actors may cope, but may also use coping to their strategic advantage.

Coping may be a straightforward consequence of a sudden increase in the demand for services. In welfare administration since 2008, officials have had to contend with a near-doubling in the number of claimants of Jobseekers Allowance (JSA) in the period of one year. The UK National Audit Office reported that officials had to radically reduce the average time spent with new claimants in order to cope with the increase, dropping way below the statutory requirement of two hours per claimant (NAO, 2013b, p14). Coping mechanisms may also become normalized in the standard practices of policy systems. In his report on UK border control, John Vine (2012, p14) found that suspensions of security checks on
passengers entering the UK had become a broadly accepted part of standard operating procedure for officials coping with sudden build-ups of queues at major UK ports. Senior officials interviewed in the border system pointed out that these checks had become legitimated as part of a culture of ‘making the system work’, often on grounds of health and safety, or in the case of French ports, responding to requests from the French authorities to deal with excessive build-up of traffic leading up to the ports.

A strong factor during the last three decades in allowing systems to rationalize and cope with demand-side pressures has been the impact of managerialism (Bastow, 2012b and 2013). The emergence of performance management and target cultures has helped policy systems to construct ‘acceptable’ levels of performance, and demonstrate success or failure against these thresholds. The dynamics here however seem dialectical. On the one hand, performance-based management has provided a sense of legitimacy that services are meeting targets in acceptable ways. For example, throughout the late 1990s and 2000s, the prison system grew a considerable number of performance targets for a wide range of its activities, including reduction of escapes, control of crowding, and provision of prisoner education and treatment programmes. On the basis of many of these targets, it is possible to argue that the system has improved its performance markedly over the last twenty years in some important aspects. Performance management has also contributed in part to helping the system to cope with the increase in prisoner numbers, and provided a basis to introduce ‘compensating’ reductions in service quality such as redefining crowding limits, cutting back on time out of cell for prisoners, and ‘disinvesting’ across the system (Bastow, 2013). However, constructed target regimes have provided a basis for distracting attention away from other policy goals. While it has provided a particular type of basis for legitimacy, it has also deprioritised focus on rehabilitative goals of prison, its formal measurement and evaluation. Up until 2010, throughout twenty years of NPM target cultures, the prison system had never had a specific target for reducing reoffending of those in its custody.
c) **Difficulties of implementing value-based goals due to resistance from within the system**

Very few, if any, organizations or systems can be seen as ‘pure’ coping archetypes. Many policy systems show clear signs of coping characteristics, but these are often countered by bouts of proactive reform. In many UK policy systems, we can identify varying degrees of improvement in performance over the decades, largely the result of actors at all levels of the system from high politics down to ‘street-level’ staff running these systems in more cost-effective and professional ways. We cannot discount here the improving effects, controversial though this will sound, of managerialism, technology modernization, outsourcing, and professionalization. Nevertheless, attempts at reform must often contend with direct or benign forms of resistance from other actors in the system. Meyer and Zucker (1989) make this very observation. ‘As power accrues to dependent actors’ they argue, ‘performance deteriorates while persistence is enhanced. Less precisely but more poignantly, politics degrades performance but enhances persistence in these organizations’ (p153).

In prisons, dependent actors such as unions have been able to wield considerable political power. Over the decades, the Prison Officers Association (POA) have influenced (indeed often held to ransom) the decisions of ministers and senior officials, particularly in constraining growth of markets for privately-managed prisons and negotiating comparatively strong pay and conditions packages. In border control, there are also signs of strongly embedded operational cultures and professional groups of frontline staff, but there is perhaps less sign of the explicitly resistant union culture that has been found in the prison system over the years. In defence procurement too, ‘top brass’ in armed forces here also play a strong part in constraining the ability of the ministers and officials to shape the overall capability to more directly suit the changing face of military challenges in a post-Cold War environment.

Resistance may also come from above in the form of veto to certain types of change that threaten to upset politically manageable equilibrium. Ministers, for example, may veto what might seem quite sensible managerial modernization. This cuts to the heart of well-established debates and questions around the relationship...
between ministers and their senior operational officials, and the extent to which these officials should or can have autonomy to manage their organizations free from excessive political influence (Page, 1992; Horn, 1995). There is a fine balance here between keeping ministers happy, and introducing managerial changes to improve value for money. This is illustrated by the controversies around suspensions of biometric checks in the border system (Bastow, 2012a). At the operational level, managers had retained the discretion to suspend biometric fingerprint checks on incoming passengers during times of heightened queue pressure at major UK airports. In April 2011 however, the Home Secretary banned the suspension of these biometric fingerprint checks on the grounds that they were undermining border security, and on finding out that they had continued between May and November 2011 (albeit at a reduced rate), she forced the resignation of senior border staff and brought the control of the border force back into the fold of the Home Office. Prior to political intervention, border force managers had been running entirely sensible risk-based pilots to introduce discretionay suspensions of checking for certain low-risk groups of passengers. The result of this controversy was that all risk-based measures were cancelled, and border officials were required to reinstate 100 per cent checks on all incoming passengers.

A more benign but equally important type of resistance is a basic function of a coping system working continually at its limits – or at least perceiving itself to do so. In coping organizations the perception is that the system is running at the peak of its capacity, regardless of whether it is or not, and hence those working inside will tend to be benignly resistant simply due to their constrained ability to take on new initiatives for change or reform. They are not resistant out of self-interest per se, but rather because the constraints of the system allows little else. The prison system, for example, has had to find practical ways of coping with a sharp rise in the number of short-sentence prisoners, and has done so by continually running its estate at levels of capacity close to 100 per cent. In doing so, it has had to constantly move the short-sentence prison population around the estate to smooth demand and fill spare capacity in order to receive new intakes. This is a complex logistical operation that has been run centrally since the mid-1990s, and has been essential ‘life support’ for coping with a crowded prison estate. The fact that these short-sentence prisoners
have to be moved around constantly in this way however perpetuates an idea that there is a limit to what extent the system can do anything with them in terms of reducing likelihood of reoffending. Actors in the system are able to recognize this as a constraint on their ability to achieve more ambitious rehabilitative goals, but it is a practical reality of the system that its stability relies on this constant disruptive movement.

d) Organisational design of the system lags behind changes in the external environment

The fourth aspect of the CCS schema is an important one. As discussed above, the concept of stress in modern policy systems suggests a problem of chronic under-supply. The question that is raised here is how capacity stress can be purely a problem of under-supply when we see many striking examples of these systems incorporating sustained obsolescence and inefficiency in the way they operate. There is a paradox here, in that actors in these systems may complain about continual financial and supply-side pressures, but they also appear to be working in systems that incorporate often striking levels of obsolescent technology, archaic standards, relative over-supply of staff, latent capacity in the way current resources are deployed, maintenance of archaic information systems, and general inefficiency in the way operations are put together.

In systems that incorporate resistance and coping characteristics, we should expect to see some degree of obsolescence and inefficiency creeping in. In the employment benefits system, for example, the ‘legacy’ effects of having to manage archaic information systems have greatly increased the costs of modernization, and hence the administration of employment benefits has remained in a perpetual state of only partly-functioning transition to a digital-era system (Dunleavy and Carrera, 2013). These legacy systems are however ‘business-critical’ in that they play a vital part in the on-going administration of benefits, and hence become sustained as an increasingly archaic part of the day-to-day operations. The syndrome here is that managerialist reforms may continually shape and squeeze cost efficiencies out of the system, but in doing so, they also entrench obsolescent ways of working, so that organizations exhibit a continual Crozerian tendency to stay significantly behind the
curve of wider society practices and technologies. Again, staff can recognize the eccentricities or inadequacies of these systems, but also make the point that they are critical to the day-to-day business of the organization. The paradox is that they are the symptoms of coping, yet are also critical to the coping process.

Similar obsolescence is conspicuously visible in the prison system (Bastow, 2013). For example, up until 2010, the main prisoner database system had been an ageing DOS system, resembling something from the mid-1980s rather than what we might expect from a modern logistic operation in the late 2000s. Prior to 2010 therefore, managers and staff were not able to get real-time information about which prisoners were in which prisons at any one time. Whereas most large logistics companies and hotel chains by the late 2000s were running their operations using sophisticated software programs and tracking devices, NOMS officials were still operating their logistical movement of prisoners around the estate with telephones, paper and pencil. Of course, senior officials interviewed explained that the inherent complexity of the prison population management was such that it had made computerization extremely difficult even in the late 2000s. Fatalism so deeply entrenched that it appears as a legitimate reason for non-change.

**Chronic capacity stress as a function of governance equilibrium**

An endless series of natural equilibria which are constantly changing as pressures change slightly. It’s like trying to design a bit of geology. The forces are greater than you can influence.

*Former senior Prison Service official in interview*

If there are blind spots in theoretical understanding of how policy systems sustain capacity stress, how can CCS help to shed light in this respect? What does it allow us to see that we would not otherwise? This assumes for a start that these blind spots contain dynamics that are worth seeing for researchers, for if they contain nothing much of interest, little is lost if we do not explore them. But the fact that many policy
systems appear to sustain symptoms of capacity stress over long periods of time should make the question interesting.

In the four potential misalignments, I argue that we have the foundations of a heuristic that can bring these blind spots usefully into more focus. They are the constituent parts of CCS, but it is in the dialectical dynamics between them, and their incorporated equilibriums, that CCS provides insights that more deterministic theoretical approaches cannot completely capture. In many policy systems, not least the few that I have mentioned here, we can fairly easily glimpse fleeting signs of capacity stress, and pulling on these threads can unravel these apparently dialectical dynamics that lie beneath. As the quote above suggests, there are ever-shifting equilibria in the governance of these systems, and understanding how to diagnose dysfunction, or to neutralise chronic characteristics, requires that we look at holistically at the aggravating and compensating dynamics within. In this last section, I discuss some implications of CCS for each of the blind spots.

First, the CCS schema shows how under-supply and over-supply dynamics can interact in important dialectical ways. In the prison system, for example, senior officials have successfully managed steadily to squeeze capacity in the estate in order accommodate population increase. Indeed, in interviews managers talk proudly about optimizing value for money by running the system continually at close to 100 per cent capacity (Bastow, 2013b). A direct consequence of this capacity tightening however is an ever-increasing reliance on administrative and logistical systems in place, and the associated difficulties of finding sufficient respite to modernize processes involved. If these processes are archaic or obsolete, when compared to wider practices in society, then further capacity tightening will likely lead to further reliance on these obsolete processes. We can use the analogy here of deploying workforce to dig ditches in a flooded field. We can increasingly squeeze the capacity of the workforce to dig faster in order to relieve the flooding, but continually doing so will only serve to increase reliance on essentially archaic processes if it continues to rain.

Similarly dialectical relationships between marginal under-supply and over-supply can be found in other sectors. The problem of sustained obsolescence from the many disjoined ‘legacy’ database systems in the Department for Work and
Pensions has aggravated the need for administrative staff to cope by normalizing often eccentric and long-winded methods for processing applications and cases for different benefits. Indeed, contractors complain about the unforeseen costs of having to find ever more permanent ‘temporary’ fixes to keep legacy systems going while new digital systems are introduced. The outcome is that Department normalizes comparatively costly and labour intensive systems. Meanwhile, a sudden increase in the number of people seeking unemployment benefits requires the system is able to double its capacity in a short period of time. Because the system is labour intensive however, it is not easily able to absorb sudden demand increase, and the only response is to increase staff numbers to scale up the processing capacity. The implications of this are double-edged. On the one hand, the Department must adapt quickly to squeeze existing capacity to absorb this increase in demand. Indeed, very few organizations, even firms, find themselves in a situation of having to respond to a doubling of demand for their services in a short time. On the other hand, the system must inevitably double the existing eccentric and long-winded processes and the inherently human resources required to implement them. Of course, budget constraints mean that staff increases are marginally less than the actual increase in demand, and hence capacity stress is tightened. Further pressure has been put on the system through the introduction of Universal Credit, thus accentuating coping, benign resistance, and obsolescence in the system (NAO, 2013a).

The second potential blind spot between strategic actor and cognitive-cultural dynamics relates to an enduring axiom of social science, and obviously CCS cannot resolve this dichotomy. The schema merely shows how understanding chronic syndromes, in whatever configuration they take, does require a focus on how individual actors at all levels of the system (from high politics through to street-level officials) assess and adapt to the real-life predicaments they face. Key to this is the interaction between self-interest and dominant cognitive norms. For example, in coping cultures, norms of finding ways to ‘make the system work’ are strong enough that they can legitimate decisions that, by the admission of these actors themselves, are far from ideal and in many cases against their own preferences or self-interest. Duty managers responsible for regulating queues at Heathrow, for example, must find ways of making the system work in terms of scaling down security measures
against in order to get passengers through immigration as effectively as possible.
This may not be ideal, but it may present an acceptable when evaluated against
political risk of ‘queue crisis’ stories breaking in the tabloid press.

If however these discretionary ‘street-level’ measures for suspending certain
types of security checks have become normalized over the years as part of coping
cultures, these norms can be used quite strategically by politicians and managers to
justify decisions taken. Arguments along the lines of ‘this is how we’ve always done
things’ can be difficult for managers or politicians to challenge, particularly in
potentially sensitive political or operational areas that can be easily destabilized (i.e.
borders or prisons). In sectors such as borders and prisons, with strong entrenched
street-level cultures and ‘dependent’ groups, it is understandable that managers and
politicians may be cautious about introducing radical change into a system that it is
in coping mode. Indeed, this links to the discussion above on how increasingly
archaic processes can creep into systems over time. In borders, this discretionary
flexibility to relax certain security checks has meant that politicians and senior
officials have collectively deprioritized the development of more sophisticated
planning and modelling tools that allow better matching of staff deployment and
fluctuations in demand at the borders. Evidence suggests that border control is still
set up in reactive mode to deal with these (predictable) fluctuations in demand (Vine,
2013a, 2013b).

But as the CCS models depicts, systems rarely exist in pure coping mode, and
actors within continually introduce or respond to bouts of reform and modernization.
In the borders system, successive governments have invested considerable funds,
particularly in the post-911 era, in modernizing the technology of the immigration
and borders system (Dunleavy et al., 2006). Managers and border staff have had to
integrate new biometric and database technologies into what has traditionally been a
fairly manual labour-intensive process. Yet contractual and implementation
difficulties have meant that these modernization attempts have delivered only partial
functionality (Vine, 2012). Indeed, the convergence of only partly functional
biometric systems and surges in passenger numbers led to airport queue ‘crises’
during the summer of 2011 and spring 2012, and hence even heavier reliance on
coping measures such as discretionary suspension. For staff and unions, stress is
further aggravated by radical reductions in resources and headcount on the frontline, again feeding into benign resistance to further modernisation (NAO, 2012).

Once we dig down to the individuals and actors facing these predicaments, we are reminded of the systemic limits to instrumental rationality. As Roberts and others have argued, achieving complete and sustained optimality in alignment of incentives is near-on impossible in ever-changing environments. For them, ‘strategic’ actors must recognize that optimality resides in sustaining ‘near-enough’ sub-optimality over time. Yet even this tends to imbue strategic actors with sufficient perspective that allows them to see the system in this broader abstract context from above. The trouble is that actors working in real-life policy systems do not enjoy such vantage point, for they are too busy influencing, coping, and resisting in dynamic ways with real-life predicaments. Seen from a great height, the concept of optimal sub-optimality makes sense to the economist. Seen from the perspective of actors working in situations of constrained autonomy, it is much harder to imagine that they can coordinate the effects of their own decisions and behaviour sufficiently, even to this extent.

Finally, what are the implications for the dichotomy between deterministic and holistic explanations? As I have tried to illustrate, certain holistic dimensions need to be present in order for us to see capacity stress in its entirety. Conceptually, the schema takes into account dynamics of reform, coping and resistance in systems, and how these interact with their external environment. In this sense, CCS belongs in a rich and well-established tradition of social science holism. Theoretically too, the schema is holistic in that it incorporates some well-established theoretical axioms. And methodologically, it is inevitable that we must combine deductive and inductive research techniques in a holistic way in order to get to this point of development of the overall concept.

But the frustrating part is that we should aspire to identify specific causal relationships. If we were to ask ‘does this CCS schema constitute a theory or a theoretical model?’ the answer would have to be no. It cannot predict a certain set of outcomes given a specific set of inputs, and this for many political scientists will be a source of frustration. Some theories have undoubtedly been important starting points in development of the schema, but I have tried to show that none of these are alone
sufficient to capture capacity stress in its entirety. There is of course scope to explore specific aspects of the CCS model in more reductionist theoretical detail, perhaps test the limitations of specific theories, and develop indicators or metrics that are able to substantiate these individual elements. It would be possible to develop indexes of ‘coping and crisis’ patterns, or reform and resistance, and begin to work these up in more sophisticated models of qualitative comparative analysis over longer periods of time. The fact also that CCS has developed out of one in-depth study of one system raises methodological questions about how to convert this form of deeper system analysis into a more systematic comparative analysis of different systems. Indeed, further questions remain about whether capacity stress is only identifiable in UK systems or whether we find similar signs of the dynamics discussed here in other countries and jurisdictions. I would estimate that we find similar examples of capacity stress outside of the UK, even though the underlying reasons and dynamics are likely to vary across different institutional, cultural, political or legal settings. CCS is a heuristic that is entirely transferable as a means to shedding light on these similarities and variations.

**Conclusions**

The analogy of pulling on a thread and watching a tapestry unravel seems apposite in describing the relationship between signs of capacity stress in policy systems, and the wider governance dynamics and operations of those systems. Signs of capacity stress come in many shapes and sizes, and as I have argued here, they are recognizable across different UK policy sectors. In diagnosing different signs of stress, it is likely that we will find a way into the CCS heuristic at a particular point, in other words, at any of its four main key misalignments:

- Value-based goals out of kilter with environment change;
- Cultures of coping and adaptation;
- Benign or explicit resistance to reform; and
- Technical obsolescence and inefficiencies in the system.
Once we enter this heuristic, it is often the case the these isolated signs of stress will interact in aggravating and compensating ways with other aspects, and these will be incorporated and incubated in the prevailing balance of the system over time.

Of course, stress is often not isolated to one system, and can be contagious in a variety of ways. Pulling on one thread in one is likely to lead us to other systems as well. For example, we might focus on the lack of cell capacity in local UK police forces to deal with those taken into custody overnight. In recent years, police forces have had to accommodate many low-level offenders who suffer from mental health illness, simply because there has been insufficient capacity in NHS mental health establishments (BBC, 2013). For the actors involved, there are often similar countervailing dynamics involved. Having police forces provide a front-line for mental patient problems in society takes costs and pressures away from actors responsible for providing mental health services. Meanwhile, police chiefs see the benefits of sustaining the perception that this surplus demand is creating a dangerous burden on local police forces and one that is unsustainable in the long term. Equilibrium therefore is dynamic across systems as well as within.

I have tried to show that setting up theoretical insights and axioms in competing ways can bring us only so far in helping to see CCS in its entirety. The CCS schema has been shaped deductively from holistic social theory, and the four misalignments have been developed more inductively through in-depth empirical analysis of the dynamics of policy systems under observable stress. Furthermore, I have set up the theoretical discussion in terms of potential ‘blind spots’ in order to show how complex systems can incorporate qualitatively distinct aspects of dichotomous pairs. It is hazardous to assume that these pairs will exist at either end of linear spectrums, as opposed to aggravating, compensating, or inherently dialectical relationships. Finally, it is in the relationships between the four aspects of the CCS schema that we find dynamic equilibriums over time, and it is the particular signs of capacity stress that must therefore be seen as characteristics sustained by these whole system equilibriums. Broadly speaking, stress can be seen as a function of these governance equilibriums over time.
Bibliography


Cairney, P. 2013b. 'What is evolutionary theory and how does it inform policy studies?' Policy and Politics, 41, 279-298.


Dunsire, A. 1990. 'Holistic Governance', *Public Policy and Administration*, 5, 4-19.


NAO 2013b. 'Responding to change in jobcentres'. London: HM Stationery Office.


Vine, J. 2012. 'An Investigation into Border Security Checks by the Independent


---

\[i\] See Easton (1967) p24

\[ii\] See Byers (2011) p15

\[iii\] King et al. (1994) point out that ‘parsimony is a judgement, or even assumption, about the nature of the world: it is assumed to be simple. The principle of choosing theories that imply a simple world is a rule that clearly applies in situations where there is a high degree of certainty that the world is indeed simple. Scholars in physics seem to find parsimony appropriate, but those in biology often think of it as absurd. [...] (p20)

\[iv\] Cairney (2013a) writes that evolutionary approaches ‘describe familiar processes of rule-influenced actions, actors adapting to their environments, and actors seeking to change the rules and their environments, and actors seeking to change the rules and their environments - in other words, the concept of ’dual causation' when 'agents interact and co-evolve with their environment'. (p290)

\[v\] See also Cyert and March (2001) on this inherent sub-optimality. They write 'we have suggested that a business firm is constrained by the uncertainty of its environment, the problems of maintaining a viable coalition, and the limitations on its capacity as a system for assembling, storing and utilizing information. As a result the theory outlined in this volume characterizes the firm as an adaptively rational system rather than an omnisciently rational system’. p117

\[vi\] Amann (2013) writes 'what is now needed, in my opinion, is a national research programme which looks critically at how a broad sample of major public public service providers in Britain actually operate, and does so with the independence and intellectual depth demonstrated, for example, by Heclo and Wildavsky almost 40 years ago in their classic study of the UK Treasury (The Private Government of Public Money)'.

\[vii\] The number of Jobseekers Allowance (JSA) claimants increased from around 743,000 to 1,352,000 between 2008 and 2009. Since 2001, the number of claimants had been stable at around 850,000.

\[viii\] Vine found that checks against criminal warnings index had been suspended around 350 times across five French ports between July 2007 and October 2011 (2012, p23). Furthermore, biometric checks had been suspended at the ‘busiest times of day at the largest airport in the UK’ (Heathrow) (p12).

\[ix\] ‘Disinvestment’ is the strongly euphemistic term to describe taking funding out of the system to establish lowest common denominator standards of service. The idea is to scale back performance that is considered to be above this standard.

\[x\] In his 2012 evaluation of these relationships in the border control system, John Vine points out the need for the ‘Border Agency's operational autonomy from the Home Office needs to be explicit [...] with a clear delineation of roles and responsibilities (2012, p9). This is something which has also been seen as a problem in the prison system in the nature of the relationship between the Home Secretary and the director general of the Prison Service over the years (Lewis, 1997).