Special Issue Article: Policy Failure

Public Policy and Administration
2015, Vol. 30(3–4) 320–341

DOI: 10.1177/0952076715581089
ppa.sagepub.com
SAGE

How to see failure: Attempts by the Inspectorate of Education to detect and disarm failure in Dutch education policy

Martijn van der Steen
NSOB, The Netherlands

Jorren Scherpenisse
NSOB, The Netherlands

Mark van Twist
Erasmus University, The Netherlands

Paul 't Hart
Utrecht University, The Netherlands

Abstract
This article looks at policy failure in systems that rely on highly autonomous organizations to deliver the services promised in the policy programs. We argue that in order to better understand success and failure in such systems, the existing categories of policy failure and organizational failure do not suffice. Therefore, we look at the ability of agents assigned to that task to effectively act in the relation in-between the system level and the level of individual organizations. This brings to the fore another type of failure – governance failure – the inability of a policy system to timely detect and asses looming failure in individual parts. It is helpful to apply a lens of interactive complexity to these systems and to look at causal loops rather than causal lines. We apply this perspective to a case of the Inspectorate of Education, to study how it dealt with three cases of looming failure in schools. The perspective of causal loops helped the Inspectorate to understand in hindsight how two schools collapsed, and it was then used proactively to intervene in a third school. The paper helps practitioners to better deal with failure

Corresponding author:
Martijn van der Steen, NSOB Lange Voorhout 17, The Hague 2514 EB, The Netherlands.
Email: steen@nsob.nl

in layered policy and sets an agenda for further research into the use of circular dynamics for policy analysis and intervention.

Keywords
Administrative theory, circular causality, complexity theory, policymaking, public administration, public management

Introduction: Success and failure in layered policy systems

In many areas of public policy, government articulates ambitious aims but largely relies on devolved, (semi-)autonomous bodies to deliver them. Government *plans*, *regulates* and *funds* to achieve policy goals, but the success or failure of it is ultimately determined by the acts of, for instance, schools, hospitals, and prisons (e.g. Clarke, 2008; Lascoumes and Le Galès, 2007; Osborne, 2006). The success and failure of policy relies on the ability of autonomous organizations to deliver what the policy system promises.

In such a layered system, the essence of a good functioning policy is not only to provide proper funding and formulate viable priorities but also to effectively manoeuvre the space between the authority of the central government and the individual autonomous organizations in the system. Policy success and failure then becomes a matter of *meta-governance*. It is the ability to keep the constitutive autonomous parts of the policy system within the boundaries of policy goals, without the ability to directly instruct and control them to do so. The success or failure of a policy system is not simply the overall and/or average systemic output but is defined by the policy makers' ability to adequately maintain and steer the mutual relation with the autonomous entities that constitute the system (Giddens, 1979). Success or failure is dependent on the ability to effectively operate *in-between* the levels of policy and delivery.

This is different from rephrasing accepted concepts of *execution* or *implementation* of policy (Hood, 1974; Pressman and Wildavsky, 1984). In many policy domains success and failure of policy is dependent on how highly autonomous organizations perform and deliver what policy programs promise. Such organizations are not the extension piece of a policy system, nor are they the long arm of the Minister responsible for the service; they are autonomous entities that operate within a broad set of boundaries that provide a mandate for a highly diverse set of practices. The absence of direct control and the mutual dependence of each other is what characterize these systems; success and failure come from effectively or ineffectively operating within those dilemmatic characteristics.

Moreover, for this type of failure it is hard to make the distinction between all-out programmatic failure and individual organizational failure. This becomes apparent when critical incidents or all-out policy fiascos occur. Success and failure can be seen as the overall performance of the system but is also assessed by the occurrence of negative critical incidents of fiascoes (Bovens and 't Hart, 1996;

Bovens et al., 2001; Gray and 't Hart, 1998; McConnell, 2010a, 2010b; Marsh and McConnell, 2010; Newman and Head, 2015; Schuck, 2014). Incidents of organizational failure hardly lower the overall quality of the system but do considerable damage to individual clients or stakeholders. Incidents damage the trust of the general public in the entire policy system and often produce costs for others in the system. Success is only partly a general metric that should on average be met; it is also to be met for each individual client. Political representatives are accountable for individual problems in individual parts of the system. For policy systems that rely on highly autonomous bodies to deliver performance, this brings to the fore the ability at the system level to effectively avert critical incidents as early as possible.

This is not to say that 'good systems' never experience organizational failure. It is evident that complex systems that deliver ambiguous services (e.g. education, health care, hold prisoners, host and/or expel asylum seekers) cannot deliver good quality every time for every single client. Things go wrong. Failure is bound to happen as a natural by-product of the dynamics in complex stems (Hogwood and Peters, 1983; Perrow, 2001; Tenner, 1997). Therefore, an essential ability of a policy system is to *detect* and *disarm* emerging instances of organization failure. This includes the ability to discern individual incidents from failure with systemic consequences (Howlett, 2012, 2014). Public organizations can have a bad year or experience difficult circumstances that make them perform badly. Most will recover by themselves and are helped best by letting them be. However, for others bad performance signals worse to come. The ability to know and see the difference and to intervene accordingly is what governance in layered policy systems is all about (Barroso, 2000; Lindblad et al., 2002; Ozga, 2009; Teisman et al., 2009; Teisman and Klijn, 2008).

In this paper, we explore how this role of detecting and disarming imminent failure in layered policy system takes place; we will look at three instances of failure within the education system in The Netherlands. Each of them was perceived in the political and public debate as signs of the inability of the system to govern the complex relation between systemic responsibilities and the performance of autonomous organizations. The Minister was directly held accountable for them and one of the cases caused very serious political problems. These instances are typical examples of governance failure. We look at how the system attempted to timely detect and disarm looming failure in three schools, to avert failure from happening. That introduces another layer to the already rich literature on policy failure and organizational failure; *governance failure*, or the inability of a policy system for metagovernance of autonomous organizations that deliver the value promised in the policy program.

The concept of governance failure fits the particular context of policy systems in The Netherlands well but it is significant beyond that; many systems in other countries and contexts are build upon the same core principle of a dual relation between a central policy core that sets goals and provides funding and autonomous parts that deliver the actual value (Barroso, 2000; Cole and John, 2001;

Lindbladet al., 2002). Therefore, insight into the dilemmas of meta-governance of failure helps to further an agenda for research and to improve the practical workings of organizations assigned to that task.

The outline of this paper is as follows. First, we start with an explanation of the perspective of our analysis of *governance failure*. Second, we discuss the methodology for our study and reflect on our position as researchers in the different cases. After that we present our findings and summarize these into general observations. In the final part of the paper, we draw conclusions and discuss the implications of our findings for academic and practical discussion about failure in layered policy systems.

Theoretical perspective: The governance of imminent failure

This paper builds on existing literature about policy failure but within it focuses on two elements. Firstly, we reflect on the relation between the policy system and the highly autonomous entities that deliver the value that policy promises. Secondly, we argue that this type of relation is inherently complex and ambiguous, and it is hence best understood from a perspective of interactive complexity. Both additions are relevant, because – as many others have already noted (e.g. Bovens and 't Hart, 1996; Bovenset al., 2001; Dunleavy, 1995; Gray and 't Hart, 1998; McConnell, 2010a, 2010b) – the literature on policy success and failure overwhelmingly draws on dichotomous understandings of policymaking and implementation, and on reductionist concepts of causality to relate effects to policy interventions. Instead of such crisp distinctions between policy and implementation, and between system and organization, more attention should be paid to the intricacies of layered governance systems and the circular, complex systems dynamics of intervening in mass public service delivery systems. They are already widespread in the delivery of government services and it is time to study more systematically how they work.

Failure 'in-between' policy and organization: Governance failure

The literature on failure can be basically divided into two bodies: policy failure and organization failure. Policy failure (Bovens, 1995; Bovens and 't Hart, 1996; Bovenset al., 2001; Brandstrom and Kuipers, 2003; Gray and 't Hart, 1998; McConnell, 2010a, 2010b) involves negative effects that can be directly related to the policy itself. Policy failure can still mean that implementation of the policy in decentred entities is the root of the problem, but that problem is assigned to the policy program itself; e.g. bad implementation due to insufficient funding, conflicting rules that make proper execution impossible, too little time to prepare properly for the execution of policy, an incomplete policy theory that does not fit the 'real' conditions of clients. Policy failure is failure *because of* the policy, in spite of how individual executive organizations or other actors perform.

Organization failure (Amburgey et al., 1993; Andrews and Boyne, 2008; Bozeman, 2011a, 2011b; Clarke and Perrow, 1996; Deeds and Pattillo, 2014;

Jas and Skelcher, 2005; Mellahi, 2005; Mellahi and Wilkinson, 2004; Sheppard and Chowdhury, 2005; Walshe et al., 2004) is failure within individual organizations that are seen as a closed system that operates within a context of policy and other external conditions, but for which the negative effects – from bad results to entire 'implosion' of the organization – are primarily the consequences of acts within that local system; e.g. bad management, internal rules not followed, groupthink in the board, the collapse of ethical standards, or mistakes in the strategic analysis of external conditions (Gruca and Nath, 1994; Janis, 1972; Jas and Skelcher, 2005; McKinley et al., 2014; Schuck, 2014). Organizational failure is about organizations that fail *in spite of* or *without* a clear connection to the policy that they are a part of.

However, there lies a gap between the two. In order to be successful – or fail – a layered policy system relies on highly autonomous semi-public organizations to deliver the policy program. In such a system, policy and organization are different entities but their acts and consequences are highly intertwined. That reduces the productivity of the distinction between the policy system and organizations that deliver services; we can crisply define the system and individual organizations, but 'quality' originates somewhere in-between.

That is more than to say that there is policy and implementation, or to construct different phases in the policy cycle for which different levels and actors in the system are responsible (Pressman and Wildavsky, 1984; Sabatier and Mazmanian, 1980; Weatherley and Lipsky, 1977); nor is it a matter of non-compliance to regulation by agents (Lodge and Wegrich, 2012; Majone, 1997). Autonomous organizations have a formalized and built-in range of freedom to decide what to do and how to do it that does not fit these dichotomies. In the case of mixed systems formulation and implementation, policy and delivery, regulation and compliance, are mutually constituted; the one takes place within the other. Organizations that deliver public goods cannot be separated from the policy, as the policy itself takes place within those organizations. That calls for reflection on how we conceptualize success and failure within these systems.

Failure as a relational concept: Governance failure

In the context of layered policy systems, rather than looking at one of the two ends of the dichotomy (organization or policy, policy or implementation, policy or compliance), it is more productive to look at success or failure at the level of the *relation* between the two. Then, success or failure is not an outcome but rather *a range of activities*; the ability of the policy core (e.g. a Ministry) to steer autonomous organizations (e.g. schools, hospitals, universities) into the intended direction of policy, without the actual instruments or hierarchical powers to simply instruct them to do so. This involves the ability to filter out looming failure and to intervene timely; not to prevent every incident from happening, but to avert cases of failure that will matter at the system level. Without the ability of a clear-cut indicator of what is what exactly, nor knowing in advance if and how an individual incident damages the system. Hence, governance failure (or success) involves the ability to monitor what

goes on in the system, direct individual schools in the intended direction, intervene as possible when deviations are detected, and fill the gap between the realm of policy and the worlds of individual autonomous organizations. Policy success and failure are far less a matter of a sound policy theory and well-assigned authorities and allocated means, but more the art of operating effectively in the *complex relations in the system*; It is the ability or inability of the responsible actors for the system as a whole to detect and disarm the imminent failure in an important (Dennard et al., 2008; Edelenbos et al., 2010; Gerrits, 2010, 2012; Jessop, 2003; Klijn, 2008). This is already done in more specialized studies of effective interventions in mixed systems, for instance in educational governance studies (e.g. Barroso, 2000; Clarke, 2008; Cole and John, 2001; Lindblad et al., 2002).

Governance failure as interactive complexity: Loops not lines

Failure is inherently a *story of causation*. In spite of a wide literature on feedback dynamics, including classic works such as those of Allison (1971), Durkheim (1997/1893), Heclo (1976), Hogwood and Peters (1983), Marx (1973), and Merton (1936), this more dynamic literature on causation remains niche in a traditional praxis (as argued by Ayres (2014), Cavana and Mares (2004), Collander and Kupers (2014), and De Roo et al. (2012)). Also, the literature on unintended consequences remains fixed on linear causal relations as the default option for understanding policy outcomes (see Ayres, 2014; Van der Steen et al., 2013). And with good reasons; for stable and bounded systems the linear causal model is a suitable way to establish a relation between cause and effect, and assign success and failure to acts of actors (Collander and Kupers, 2014). However, when applied to unstable or complex systems the linear perspective presents two shortcomings.

Firstly, it does not take into account interactive dynamics. Interventions do not stop at the designated target or time period but continue beyond. The linear perspective assumes the causal effect of A to be *bounded* to B. However, in complex systems it is difficult to project beforehand where effects will 'go', how long they resonate, and who will respond to it. What is called an 'unexpected outcome' from the linear view on causality (Sieber, 1981: 10) can be *expected* from the perspective of interactive complexity. Secondly, the linear perspective on causation hardly takes into account *reflexivity* or the *learning capacity* of agents. When policy is added to system agents learn from what happens. Over time they will change their response to measures. What seemed to work the first time probably plays out differently the next because the previous intervention changed the system; it instigated learning in the system and as a result agents may behave differently next time. Causation is a dynamic process that evolves over time, rather than a fixed, stable and almost a-temporal relation between cause and effect.

To explain the occurrence of failure in contexts of interactive complexity, it is imminent to look beyond linear and fixed mechanisms of cause and effect. To see failure differently we have to look through a lens of causality that takes into account the *dynamics* in the system (Byrne and Callaghan, 2013; Edelenbos

et al., 2010; Gerrits, 2010, 2012; Pawson, 2006; Teisman et al., 2009; Teisman and Klijn, 2008; Tilley, 2010; Van der Steen et al., 2013). Therefore, we apply the lens of circular causality to the study of policy failure and look at causal loops. Circular causality originates in the system dynamics and cybernetics literature, and it is applied in the context of policy and system analysis (Cavana and Mares, 2004; Forrester, 1961; Haraldsson, 2000; Juarrero, 2011; Morçöl, 2010, 2012; Morçöl and Wachhaus, 2009; Perrow, 2001; Senge, 1990; Streeck and Thelen, 2005; Weick, 1995; Weick and Sutcliffe, 2001; Wildavsky, 1988). Central to that is the interconnectedness of elements and the feedback mechanisms that shape the interactions between them. It considers outcomes the effect of interrelated interactions between different actors and factors of the system (Forrester, 1961; Gerrits, 2012; Haraldsson, 2000; Pierson, 1993; Pilkey and Pilkey-Jarvis, 2007; Richardson, 1991; Richmond, 1993; Teisman et al., 2009; Tenner, 1997).

Actions generate feedback, which becomes input for others; the feedback loops create patterns, often in the form of *loops* (Collander and Kupers, 2014; Merali and Allen, 2011). The literature discerns different basic shapes of causal loops. Some loops are *self-balancing* (Haraldsson, 2000; Lane, 2008; Maruyama, 1963; Merali and Allen, 2011; Richardson, 1986; Senge, 1990; Toole, 2005). Others display a *self-reinforcing* pattern; a change in one factor enforces a loop that leads to a magnification of the original effect (Haraldsson, 2000; Lane, 2008; Maruyama, 1963; Merali and Allen, 2011; Richardson, 1986; Senge, 1990; Toole, 2005). Some loops draw the system more towards an outcome intended by the policy maker, while others pull it further away from what was meant (Morçöl, 2010).

Systems that are dominated by self-balancing loops have in-built mechanisms that draw towards status quo; disturbances are corrected through the self-balancing patterns (Haraldsson, 2000; Morçöl, 2010; Teisman et al., 2009). The opposite goes for systems with strong or dominant self-reinforcing loops; then, originally minor interventions can accelerate into large development in the system that can flip the balance of the system as a whole; this can be positive in terms of intended outcomes but also negative. Sometimes a system develops into a virtuous circle that generates excellence; sometimes a system is locked in a vicious cycle that eventually leads to its downfall.

Given the relational nature of governance failure the perspectives of interactive complexity – of causal loops – help to better understand policy failure in complex layered policy systems (Richardson, 1986, 1991; Richmond, 1993; Toole, 2005; Van der Steen et al., 2013). The perspective of causal loops offers a rich account of the process of failure that can help researchers and practitioners to interpret (looming) instances of failure. It helps to make sense of dramatic outcomes and supports the development of the practical repertoire for detecting and disarming looming failure.

How to 'see' failure: Loops to detect failure and to intervene effectively

In the remainder of this article, we use the combined perspective of *governance* and *interactive complexity* to look at instances of failure in an existing and well-studied

layered policy system, that of education policy in the Netherlands. We apply the lens to acts of the Inspectorate of Education, which in that system is the agent of governance that operates in the in-between space between the policy core (the Ministry) and the autonomous units of delivery (schools). This relation between system and school, and the intermediary role of the Inspectorate, is a typical example of a layered policy system and can act as an exemplary case for exploring our perspective (see Ehren et al., 2005, 2013; Gustafsson et al., 2014; Janssens, 2007; Jones and Tymms, 2014; Klerks, 2013; Wilkins, 2014). Taken together, the concepts of governance failure and interactive complexity provide a different perspective to see *failure* differently. This *different seeing* has dual meaning; it involves a different perspective to analyse failure in hindsight but also provides new tools to detect looming failure and disarm it proactively.

Research approach: A study into the relation between the Inspectorate and failing schools

Research focus: Ensuring quality of education in the Dutch education system

In this paper we examine governance failure in the education sector in the Netherlands. Our research objective is to better understand the dynamics that lead to the outcomes in these concrete cases. We look at instances of failure in three schools, to understand how governance attempts to detect and disarm imminent failure worked out.

Context: A built-in duality in the education system

In the Dutch primary, secondary, and tertiary education sector, *schools* act under the responsibility of the *Minister of Education*. They receive funding from the central government and have to achieve end terms (outputs) assigned by the Ministry. However, within that framework schools are autonomous entities as consigned in Article 23 of the Constitution. Not-for-profit groups, often of religious denominations or espousing a certain educational philosophy, own most schools. A smaller amount of schools is public and are operated by the Municipal government. These schools hold the same degree of autonomy. The Minister of Education has no authority in the administration of schools, nor can it influence the education process in the classroom.

The inspectorate as a tool of governance in the education system

The Inspectorate of the Ministry of Education operates in the space between the *policy program* of the Ministry and the *delivery of education* by autonomous schools. The Inspectorate conducts regulatory supervision of schools to check if schools produce the required education results, which are measured in standardized tests. Moreover, the Inspectorate oversees if schools comply

with baseline standards for basic facilities, e.g. quality of housing, solidity of the financial position, and the required level of training of the teaching staff. The Inspectorate ensures the overall quality of the system on behalf of the Minister.

Furthermore, the Inspectorate plans periodic inspection visits to schools. It uses data from test scores to monitor the system and to develop a 'risk image' for each school. The inspection uses this to allocate capacity for onsite inspection; weak schools are inspected more frequently and better schools are allowed so-called inspection vacations.

The risk image, along with statistics and data about schools' performance, is published on the web; this is part of the 'naming, faming, and shaming' strategy of the Inspectorate. Parents can compare schools and use exit and voice to correct schools. If schools do not comply with basic standards or fail to reach required scores, they slip into a heightened risk category. The inspectorate then intervenes in escalating steps. Firstly, schools are paid extra *inspection visits* to conduct a broader assessment of the school. Secondly, in case of repetitive underperformance schools are placed under *intensified supervision*. The Inspection now regularly visits and assesses the school. Moreover, the school gets two years access to extra means for mandatory improvements such as education materials, training of teaching staff, and assistance for the schools' management. Thirdly, if a school does not improve within two years after the start of the intensified supervision the Minister will *cease financial support* for the school. This in effect means the closure of the school, because the School Board will not be able to find enough guaranteed alternative funding for the school. Note that closing a school is not possible in the direct sense of the word. The Minister does not have formal authority over individual schools that lies with the Board of the school. Also note that in this system schools 'fix themselves' even when under intensified supervision. The Inspectorate measures how they do but will not *intervene* in the school. The Inspectorate performs a crucial role in ensuring the quality of each school in the system but does so from a position in-between the policy core and the individual schools that deliver education.

Research methods: Participative research with the Inspection

We conducted a participative research: we were invited by the Ministry of Education to participate in two projects that dealt with underperforming schools, both in a very different way. In this paper we put the two together, because in their combination they reveal interesting insights in how to see policy failure in layered policy systems. We look at three examples of failure in individual schools and see how the Inspectorate attempted to detect and disarm them. For us, these three failing schools are merely the *locus* for our *focus*; we are really interested in how the Inspectorate operates in-between the policy core and the schools and how the perspective of causal loops did or did not help the Inspectorate to detect and disarm imminent failure in the policy system.

Project 1: School A and B, ex post analysis

The first two cases – school A and B – were ex post cases of failure; the collapse of these schools was the reason for the Ministry of Education and the Inspectorate to ask for an account of what happened. Both schools had been under close scrutiny of the Inspection. For case A the feeling was that the interventions of the Inspectorate somehow had made matters worse; for case B there was an idea that Inspectorate missed crucial early warnings. The Ministry and the Inspectorate wanted to learn from these cases to avert more of them to happen, and we were called in to reconstruct what happened.

We conducted desk research to study all of the relevant documents about the school and we reconstructed the chronology of the years until the moment of its downfall. After we finalized our chronology, we conducted a series of interviews with all of the relevant stakeholders in and around the school. We asked stakeholders to reflect on the chronology, check it as a factual account, and provide us with their personal analysis of the process. We used these to construct the dynamic patterns (causal loops) we thought had led to the collapse of the school. We then organized a meeting to present our account of what had happened, present the loops, and to check if the stakeholders recognized our reconstruction. For both cases, the stakeholders accepted our account and we finalized our report. Then, the report about School B was used as 'fact base' of the Parliamentary Research Commission that investigated the collapse of School B. The report about School A was used internally by the Ministry and the Inspectorate, and was presented to the OECD Committee on improving educational quality, as lessons from the Dutch approach of dealing with underperforming schools (Van Twist et al., 2013).

Project 2: School C, 'real-time' participative observation

After we finalized the reports of School A and B the case of School C came up. School C was an Islamic secondary school that was under increased scrutiny of the Inspection. Already for years it suffered from big problems, and the Inspection and the Minister asked us as researchers to observe the team of Inspectors as they dealt with this school. School C was still operational at the time of our research. The Inspectors wanted to apply ex ante the lessons learnt ex post in School A and B. We were assigned as members of the special inspection team and went into the school with them. The purpose of this special team was to do a quick but thorough analysis of the school and its surrounding system, in order to assess the sustainability of the school. The Minister granted the team two weeks to come to its conclusion, a time frame that was exceptionally tight because of intense media attention and public outrage over incidents at the school. We went along with them as they took office in a classroom in the school and observed how the Inspectors came to their judgement about the school.

We observed the Inspectors in action. Most of the time we confined ourselves to making field notes. Towards the end of the process we participated in strategic meetings in which the decision was eventually made to terminate funding for the school. We

did not directly influence that decision but we did participate in the construction of the analysis that informed it. The Inspectorate made a deliberate attempt to use insights from the first two cases to interpret the situation and intervene in the school, and therefore asked for our assessment of the situation several times.

Combining the cases to learn about the governance of failure

The two cases are not comparative but altogether form a narrative about how the Inspectorate attempts to detect and disarm failure in the education system; for us, the Inspectorate is the real case. The interaction between the schools and the Inspectorate provides an opportunity to explore the complex reality of failure in mixed policy systems (Rhodes et al., 2007; Weiss, 1994). We do not use the case to prove a theory but reflect on a deliberate attempt by an agent in a mixed policy system to deal with failure. By doing so, we explore if our perspective can help the governance of success and failure in such systems. Moreover, we want to explore the consequences of that for practical dealing with imminent or perceived failure in mixed policy systems.

Analysis: The praxis of in-between governance of failing schools

School A: How an attempt at improvement made matters worse

School A is a small primary school located in a densely populated countryside. The school accepts all students and takes in a relatively large number of special needs children. There are two other schools in close proximity.

In the last phase before labelling it as ‘very weak’, the Inspectorate conducted a close study at School A. The Inspectorate wanted to look ‘beyond’ the test scores and analyse its strong and weak points. Inspectors observed teachers at work, talked to the management, interviewed the Board, and had meetings with parents. The Inspectors developed an understanding of the problems of the school. The school was small and ‘suffered’ from its policy to accept all special needs children; they weighed heavy in test scores. Also, the staff looked strained. The school’s management was not considered a problem. The Inspectorate expected that a strengthening of the teaching capacity of the staff would help the school recover. Inspectors appreciated that the school accepted special needs children.

After the Inspectorate assessed the school as ‘very weak’ the standard protocol went into practice. The ‘very weak’ status was published on the website of the Inspectorate, and the school automatically enrolled in a two-year improvement program. After that, the situation changed rapidly. The qualification came unexpected for parents. They took it hard that they were not informed earlier. The school was small and parents had a good contact with the director and staff. The school scrambled to inform parents in a special meeting; they were told how the school took up an improvement programme and how that would probably work. The inspectors were present at this meeting but they did not formally participate.

The meeting did not have the intended effect. Shortly after the announcement, parents of good performing children transferred their children to other schools; this tipped the balance in the school to the weaker students. More transfers followed; the parents who moved were known faces within the school, and their departure was perceived a strong signal of upcoming decline.

When the new school year begun classes had to be combined because there were not enough students for separate groups. This meant extra pressure on teachers and less time for individual students. Furthermore, staff members had to be laid off due to the decline in enrolment. Parents were worried because they saw hard working teachers being laid off. To make matters worse, the increased pressure burdened teachers. Some turned in sick, which caused problems for the staff as a whole. As one teacher told us: 'Two teachers left, we had a lot of substitutes, and then another colleague left as well because it was unbearable'. Before the improvement programme was even kicked-off the school had already dropped far below the initial assessment by the Inspectorate. Sadly, the announcement of upcoming improvement had initiated a steep decline of the school. Two years after the initial intervention by the Inspectorate the school had to be closed.

During this process the Inspectorate *observed* the downfall of the school; inspectors frequented the school to assess its progress but they were not allowed to do anything. They saw teachers struggling to use new methods. Inspectors were allowed to provide feedback. This irritated the staff. As one teacher stated:

> It is demotivating that your efforts are not seen. They [the inspectors; our addition] sit in the back of the classroom, observe what I do, but say nothing. (. . .) The only thing that matters is if my five special need children perform at the test.

During the meeting, the inspectors confirmed that they had struggled with this. They noticed that teachers improved and saw how much effort teachers put into the special needs children. They found it hard that they could hardly use their observations. They could not report signs of improvement because the protocol did not allow it. One inspector stated: 'I was sorry that I could not directly talk to parents and tell them how good the school was doing. That might have persuaded them to stay. It could have tipped the balance for the school'. Another inspector shared a similar thought: 'I think teachers wanted to hear that we appreciated their efforts. But we could not talk to them in that way'.

Our analysis of the developments in School A was that the intervention of the Inspectorate triggered an accelerating cycle of collapse. The initial announcement scared some of the most active parents and their transfer to other schools resulted in a cascading development of more transfers. Staff had to be laid off and that put more pressure on teachers. The special needs children not weighed heavier in the averages of the test scores. While most children were doing okay, the overall performance of the school was insufficient.

The Inspectors at the school recognized this pattern, beforehand and during the process. Especially the public announcement of the 'very weak' status had caused

them worries from the beginning, as they foresaw the possible chain reaction it might spur. However, the protocol did not provide space for deviation. Also, in hindsight inspectors told us they felt bad that they were not able to provide feedback to teachers. Inspectors told us they think all that might have helped.

School B: How a large school slid to bankruptcy unnoticed while under close supervision

School B is one of the largest school boards in the Netherlands, with over 30,000 students and 3000 staff delivering secondary and college level education. School B is the only school board in The Netherlands that mixes secondary and college level education. Most of its locations are in two of the biggest cities. Over the years, the Inspectorate intensified the educational quality supervision on several schools under the Board of School B and placed the Board under increased financial supervision.

For several years, school B was considered an excellent example of innovative vocational education. The school generated a lot of capital to invest in new buildings and expensive new equipment for the schools. However, problems emerged. The driver behind the model was the prognosis of *growth* of the number of students; investments were made to build the environment needed for new innovative educational concepts to work but they would only pay off with the enrolment of new students. The Board anticipated that once the school would grow, it would push all of the others out of the regional market and establish a de facto monopoly. From there growth would be self-sustaining.

However, investments laid a heavy burden on the budget for education. Management economized on staff. The Inspectorate saw how some of the schools slid down towards 'very weak' quality. Inspectors conducted onsite inspections to look at those schools; they reported that school B faced serious problems with the quality of teaching. Moreover, student dropout increased and more teachers called in sick. Students complained about a lack of personal attention. The news was picked up by local media and enrolment dropped further.

The decline in enrolment caused financial problems. In order to attract more students, the Board invested more money in buildings; the capital came from loans and the solvability worsened. In order to sustain the innovative image, the school patched up its balance, by pricing in overconfident numbers of enrolment. The financial reports of School B caught the attention of the Financial Unit of the Inspectorate. The school was placed under close Financial Supervision by the Financial Unit, which meant that the Board of School B was required to send its annual reports directly to the Financial Inspector. He would monitor the developments of the schools' key financial indicators. Also, the school was required to send a special statement by school's accountant that the Board was 'in control'. The accountant issued these reports and the schools' board reported its financial statements in time. The school remained under close financial inspection but seemed to have stabilized.

After another critical media report – this time a televised documentary – enrolment dropped further. This triggered the Board into an all or nothing strategy.

On paper the Board made promising plans in which they counted on lower interest rates and on a projection of steep growth in enrolment. In reality, the banks were reluctant to renew credits and wanted guarantees from the Ministry of Education. That was not possible, because it would reveal the severity of the situation. Meanwhile, the Ministry and the Inspectorate were bounded by their own procedures and were not allowed to look beyond the official financial figures the Board presented, all of which were officially verified by the accountant of school B.

After a strenuous negotiation with School B the banks declined a new credit, just as the Ministry had decided that based on the new prognoses by the Board – validated by the accountant – the school could be released from intensified financial supervision. The Board had to turn to the Ministry and this time it was formally required to present the entire financial administration, not just the polished version of it. In order to save the school and secure the education of the 30,000 students, the Ministry was forced to take over the debts of School B. School B was bailed out at the expense of tax money.

Our analysis of the developments in School B is that it got caught in a slow moving but self-sustaining negative cycle; it had invested heavily and in order to pay off those investments the school had to grow. However, the investments reduced the funds available for education and that decreased enrolment. Staff felt strained and reported sick. That increased financial pressure and so on and so forth. The Inspectorate had all the information to 'see' this cycle by itself but was unable to combine insights from various units within the Inspectorate into one picture of the dynamics in the school. The information was literally divided over different units in the Ministry.

By themselves neither the financial nor the quality issues were severe enough to allow the unprecedented act of intervening directly into the school and – for instance – ask for the underlying financial data or to see the internal data about teaching and quality. Had the Inspection done that, it would have seen how bad the situation actually was. Inspectors told us they had a bad feeling about the School all along, but never found a trigger to 'scale up'. In hindsight the pattern looked evident, but within the formal procedures it was difficult to discern it in advance. Also important was that the process went *slow*; there were no big disruptors or evident crisis in the school, it all went a little bit worse each year. The key was to combine information and see how different aspects of the system interrelated with each other but that was not build into the system. The system can very effectively scale up in case of a crisis but was not able to detect a slow-burning crisis building up until it eventually led to the ultimate collapse. The Inspectorate had all the dots, but the protocol did not enable inspectors to connect them.

School C: A school closed for reasons of prospective non sustainable dynamics

School C is a medium-sized secondary school. It was the first Islamic secondary school in the country. From its start 10 years ago, problems were manifold at

School C. However, each time the school managed to resolve the most acute problems. Also, in spite of its problems the school receives strong support from the local Muslim community.

School C had been under close attention of the Inspectorate. Years earlier, the school was condemned for the use of money for educational improvement for a study trip to Mecca and received a fine that caused immediate financial problems. Educational quality decreased because of that and the school was depicted as *very weak* by the inspectorate. To find new financial means the school struck a deal with the municipality to get a loan; however, the terms of the loan were very bad for the school, with high interest rates that burdened the schools for years to come.

In a critical internal review, the schools' director was signalled as an important negative factor in the school. He was replaced by an interim director who quickly took up major improvements. That helped, but the fee of the interim director was expensive. Combined with the high interest on the loan, this caused budgetary problems. Another problem was that the municipality placed the school in an outdated building; it provided sufficient space for the growing school, but the costs for maintenance and energy doubled. This was too much to bear for the already strained budget. The Inspectorate increased the financial supervision and urged the schools' director to quickly improve the solvability of the school. Again, the school managed to improve towards slightly better than the minimal requirements, partly because of the quality of the interim director.

A year later, 22 central (national) exams were stolen in the school and were sold over the Internet. This caused turmoil as the national exams had to be postponed. The Inspectorate was called and it assigned a special team to go into the school and to assess the future of the school. For the Inspectorate, this was a difficult task. It has no norms or protocols for a *prospective* judgement about a school. The Inspectorate normally relies on retrospective judgement from confrontation of test results with set standards. With the recent experiences in School A and B in mind, the Inspectorate wanted to take into account the dynamics of the school; they wanted to know why new problems for the school kept popping up, and what given those dynamics the chances of the school were.

The Inspectorate went into the school and conducted an intense review of the school to describe the causal pattern of the school. The team interviewed all relevant stakeholders and reconstructed the time line of the school from different angles and perspectives (financially, quality, relations with other actors, quality of the board). The inspectorate concluded that the different key stakeholders around the school were entrenched in positions that would continue to renew and aggravate problems after they would have been temporarily solved. According to the Inspectorate the school was caught in a self-sustaining negative cycle, in which the cumulative actions of different stakeholders would reproduce the same problems all over again. The Inspectorate concluded that the pattern in the dynamics of the school could not be turned around; because of that the only way to avert more problems would be to close the school. School C was not closed

for its current problems but for prospective reasons; today's problems did not provide reason to close the school, but the analysis of the underlying dynamics did. This was a novelty in The Netherlands, and although it remains contested the Ministry of Education and the Inspectorate see it as a possible new standard for the assessment of the quality of schools.

Conclusion

Our three cases show the interactive dynamics between the system level of education policy and the autonomous actors that deliver education. The cases show how the interactive dynamics can be understood in terms of circular dynamics; the dynamics between actors explains outcomes. Therefore, one would expect the repertoire for the governance of it – in this case by the Inspectorate – to be based on a dynamic concept of causality: it seems logical to take loops into account when looking at cases of failure in a complex system. The concept of interactive complexity puts interactions to the fore for an analysis of how systems develop and effects occur; the interactions between actors take the form of loops, which can be self-balancing or self-reinforcing; and move towards or lead away from the expected policy goals. Loops consist of interactions between the actors in the policy system, which makes this explanation essentially *local*; loops depend on the local balance of actors, local contextual factors, and local processes that direct the dynamics in a certain direction.

However, in School A and B we saw how current systems of the Inspectorate leave little room for such dynamics and follow a protocol that relies on linear causal mechanisms; the system follows a logic that leaves little room for local knowledge, even though inspectors acquire rich knowledge about the dynamics in a particular system. It is interesting to see that cases do not reveal a *lack of information*, nor a *lack of understanding of loops* by inspectors; they reveal how such insights and information are not used, because the perspective that is used does not enable inspectors to *see* the linkages. Therefore, we see these two schools as instances of *governance failure*; a failure to appropriately work in the in-between space of policy and organizations, in order to detect and disarm failure. We see this as 'failure', because the Inspectorate did possess the information it needed to act more effectively. We concur that this cannot be predicted with certainty, but it is safe to say that a circular perspective would have allowed the inspectors to act differently. Moreover, the inspectors we interviewed and who attended our feedback sessions *themselves* argued that they could have averted failure. Had they looked differently, they would have seen failure looming and would have intervened differently to disarm it. Early detection and disarming of failure in a layered policy system requires a different way of looking at and seeing failure; a circular perspective of loops, rather than a perspective of linear causation.

Therefore, it is interesting to see what the Inspectors did in School C, and why we do *not* see it as an instance of governance failure. The school was closed and it is hard to call that a success. But what did go 'right' at school C was that the

Inspectorate took up a perspective to look at failure that matched the characteristics of what they were looking at; namely a complex system, driven by interactive dynamics between different stakeholders in and around the school. When inspectors used the lens of interactive complexity to map the dynamics in the school, they quickly came to the conclusion that the school did not have a sustainable future, even though each recent incident alone was not strong or bad enough to close the school. School C was closed for prospective reasons, as a matter of prevention for worse to come. Again, this is not a normative plea for prevention, but it is an interesting act of deliberate use of a complexity perspective to intervene in a complex policy system. The analysis of the Inspectorate and the decision that followed from it were highly contested; by parents who were angry that the school was closed, but also by people, politicians, and academics who criticized the unprecedented act of closing a school for prospective reasons. However, the argument held firmly, also in Parliamentary and in the public debate about it. That is not at all evidence that the Inspectorate was right, but it does show that this way of looking at schools and failure holds promise to produce understandable and rich analyses of how and why schools fail. And it definitely calls for a more thorough debate about the consequences of applying this perspective to practical use. It flips how we organize systems and procedures, and calls for new repertoire for deliberation and intervention; because system dynamics is inherently *local* it is important to build organizations and protocols that are able to discern and value local judgment – and the variety that will inevitably arise from it. This is a very fundamental point for consideration, as it comes near to deeply embedded values such as equal treatment. Also, the use of a complexity perspective can be used proactively, to disarm failure before it happens; this is not fundamentally different from what systems currently do, but the foundation for such preliminary actions is only marginally 'objective' and relies heavily on local professional judgment and interpretation. In order for a further and more systematic application of the complexity perspective into systems of policy and governance, it is imminent that these local and professional elements are integrated into systemic structures, protocols, and procedures. This is a highly relevant frontier for both academic research and organizational practice.

Finally, it is important to come to new and better definitions of success and failure. In this article, we have argued that the understanding of failure in layered policy systems requires us to look beyond crisp distinctions between policy failure and organizational failure, and of policy formulation and implementation. For mixed systems that rely on highly autonomous organizations, to deliver policy success and failure are not defined as outcomes, they should be seen at a relational level; the ability to act in the relation between the system level and the organizational level, and from there to detect and disarm failure as soon and as good as possible. That is hardly a crisp metric for defining failure, and it raises many new questions. For now, it is enough to break open current definitions of policy success and failure in this way, even though we lack a proper definition to replace them. It is now imminent to further explore the agenda of governance failure and develop more practical concepts and

definitions to work from. We hope that this paper and this Special Issue can be first steps in that journey.

References

Allison GT (1971) *Essence of Decision: Explaining the Cuban Missile Crisis*. 3rd edition. Boston: Little Brown.

Amburgey TL, Kelly D and Barnett WP (1993) Resetting the clock: The dynamics of organizational change and failure. *Administrative Science Quarterly* 38: 51–73.

Andrews R and Boyne GA (2008) Organizational environments and public-service failure: An empirical analysis. *Environment and Planning. C, Government and Policy* 26: 788.

Ayres S (ed) (2014) *Rethinking Policy and Politics: Reflections on Contemporary Debates in Policy Studies*. Bristol: Policy Press.

Barroso J (2000) Autonomie et mode de régulation dans le système éducatif [Autonomy and the form of regulation in the education system]. *Revue Française de pédagogie* 130: 57–71.

Bovens M (1995) Frame multiplicity and policy fiascoes: Limits to explanation. *Knowledge and Policy* 8: 61–82.

Bovens M and 't Hart P (1996) *Understanding Policy Fiascoes*. New Brunswick: Transaction.

Bovens M, 't Hart P and Peters BG (2001) Success and Failure of Public Governance: A Comparative Analysis. Cheltenham: Elgar.

Bozeman B (2011a) Toward a theory of organizational implosion. *The American Review of Public Administration* 41(2): 119–140.

Bozeman B (2011b) The 2010 BP Gulf of Mexico oil spill: Implications for theory of organizational disaster. *Technology in Society* 33: 244–252.

Brandstrom A and Kuipers SL (2003) From 'normal incidents' to political crises: Understanding the selective politicization of policy failures. *Government and Opposition* 38: 279–305.

Byrne D and Callaghan G (2013) *Complexity Theory and the Social Sciences: The State of Art*. Oxon: Routledge.

Cavana RY and Mares ED (2004) Integrating critical thinking and systems thinking: From premises to causal loops. *System Dynamics Review* 20: 223–235.

Clarke J (2008) Performance paradoxes: The politics of evaluation in public services. In: Davis H and Martin S (eds) *Public Services Inspection in the UK. Research Highlights in Social Work* (50). London: Jessica Kingsley Publishers, pp. 120–134.

Clarke L and Perrow C (1996) Prosaic organizational failure. *American Behavioral Scientist* 39: 1040–1056.

Cole A and John P (2001) Governing education in England and France. *Public Policy and Administration* 16: 106–125.

Collander D and Kupers R (2014) *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*. Princeton: Princeton University Press.

Deeds V and Pattillo M (2014) Organizational "failure" and institutional pluralism: A case study of an urban school closure. *Urban Education*, pp. 1–31. Epub ahead of print 26 January 2014. DOI: 10.1177/0042085913519337.

Dennard L, Richardson K and Morçöl G (2008) *Complexity and Policy Analysis: Tools and Concepts for Designing of Robust Policies in a Complex World*. Goodyear, AZ: ISCE Publishing.

De Roo G, Van Wezemael J and Hillier J (eds) (2012) *Complexity and Planning: Systems, Assemblages and Simulations*. Farnham: Ashgate Publishing, Ltd.

Dunleavy P (1995) Policy disasters: Explaining the UK's record. *Public Policy and Administration* 10: 52–70.

Durkheim E (1997/1893) *The Division of Labour in Society*. New York: The Free Press.

Edelenbos J, van Schie N and Gerrits L (2010) Organizing interfaces between government institutions and interactive governance. *Policy Sciences* 43: 73–94.

Ehren MCM, Altrichter H, McNamara G, et al. (2013) Impact of school inspections on improvement of schools—describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability* 25: 3–43.

Ehren MCM, Leeuw FL and Scheerens J (2005) On the impact of the Dutch Educational Supervision Act: Analyzing assumptions concerning the inspection of primary education. *American Journal of Evaluation* 26: 60–76.

Forrester JW (1961) *Industrial Dynamics*. Cambridge, MA: MIT Press.

Gerrits L (2010) Public decision-making as coevolution. *Emergence: Complexity and Organization* 12: 19.

Gerrits LM (2012) *Punching Clouds. An Introduction to the Complexity of Public Decision-Making*. Litchfield Park: Emergent Publications.

Giddens A (1979) *Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis*. Berkeley: University of California Press.

Gray P and 't Hart P (1988) *Public Policy Disasters in Western Europe*. London: Routledge.

Gruca TS and Nath D (1994) Regulatory change, constraints on adaptation and organizational failure: An empirical analysis of acute care hospitals. *Strategic Management Journal* 15: 345–363.

Gustafsson JE, Lander R and Myrberg E (2014) Inspections of Swedish schools: A critical reflection on intended effects, causal mechanisms and methods. *Education Inquiry* 5: 461–479.

Haraldsson HV (2000) Introduction to systems and causal loop diagrams, System Analysis course, Lund University (January), 1–33.

Heclo H (1976) Conclusion: Policy dynamics. In: Rose R (ed.) *The Dynamics of Public Policy: A Comparative Analysis*. London: Sage Publications.

Hogwood B and Peters BG (1983) *Policy Dynamics*. Boston: Wheatsheaf Books.

Hood C (1974) Administrative diseases: Some types of dysfunctionality in administration. *Public Administration* 52: 439–454.

Howlett M (2012) The lessons of failure: Learning and blame avoidance in public policy-making. *International Political Science Review* 33(5): 539–555.

Howlett M (2014) Why are policy innovations rare and so often negative? Blame avoidance and problem denial in climate change policy-making. *Global Environmental Change* 29: 395–403.

Janis IL (1972) *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*. Oxford: Houghton Mifflin.

Janssens FJG (2007) Supervising the quality of education. In: Bottcher W and Kotthoff HG (eds) *Schulinspektion: Evaluation, Rechenschaftsleging und Qualita tsentwicklung [School Inspection: Evaluation, Accountability and Quality Development]*. Munster: Waxman.

Jas P and Skelcher C (2005) Performance decline and turnaround in public organizations: A theoretical perspective and empirical analysis. *British Journal of Management* 16: 195–210.

Jessop B (2003) Governance and meta-governance: On reflexivity, requisite variety and requisite irony. In: Bang HP (eds) *Governance as Social and Political Communication*. Manchester: Manchester United Press, pp. 101–116.

Jones K and Tymms P (2014) Ofsted's role in promoting school improvement: The mechanisms of the school inspection system in England. *Oxford Review of Education* 40: 315–330.
Juarrero A (2011) Causality and explanation. In: Allen P, Maguire S and McKelvey B (eds) *The Sage Handbook of Complexity and Management*. London: Sage, pp. 155–163.
Klerks MCJL (2013) The effect of school inspections: A systematic review. *School Improvement*. Paper presented at the ORD, Wageningen, The Netherlands, 20–22 June 2012, pp. 2–32.
Klijn EH (2008) Policy and implementation networks. Managing complex interactions. In: Cropper S, Ebers M, Huxham C, et al. (eds) *The Oxford Handbook of Inter-organizational Relations*. Oxford: Oxford University Press, pp. 118–146.
Lane DC (2008) The emergence and use of diagramming in system dynamics: A critical account. *Systems Research and Behavioral Science* 25: 3–23.
Lascoumes P and Le Galès P (2007) Introduction: Understanding public policy through its instruments – From the nature of instruments to the sociology of public policy instrumentation. *Governance* 20: 1–21.
Lindblad S, Ozga J and Zambeta E (2002) Changing forms of education governance in Europe. *European Educational Research Journal* 1: 615–624.
Lodge M and Wegrich K (2012) *Managing Regulation. Regulatory Analysis, Politics and Policy*. Hampshire: Palgrave Macmillan.
McConnell A (2010a) Policy success, policy failure and grey areas in-between. *Journal of Public Policy* 30: 345–362.
McConnell A (2010b) *Understanding Policy Success: Rethinking Public Policy*. Basingstoke: Palgrave Macmillan Ltd.
McKinley W, Latham S and Braun M (2014) Organizational decline and innovation: Turnarounds and downward spirals. *Academy of Management Review* 39: 88–110.
Majone G (1997) From the positive to the regulatory state: Causes and consequences in the mode of governance. *Journal of Public Policy* 17: 139–167.
Marsh D and McConnell A (2010) Towards a framework for establishing policy success. *Public Administration* 88: 564–583.
Maruyama M (1963) The second cybernetics: Deviation-amplifying mutual causal processes. *American Scientist* 51: 164–179.
Marx K (1973) *On Society and Social Change*. Chicago: The University of Chicago.
Mellahi K (2005) The dynamics of boards and directors in failing organizations. *Long Range Planning* 38: 261–279.
Mellahi K and Wilkinson A (2004) Organizational failure: A critique of recent research and a proposed integrative framework. *International Journal of Management* 5: 21–41.
Merali Y and Allen P (2011) Complexity and systems thinking. In: Maguire S, Allen P and McKelvey B (eds) *The Sage Handbook of Complexity and Management*. London: Sage.
Merton RK (1936) The unanticipated consequences of purposive social action. *American Sociological Review* 1: 894–904.
Morçöl G (2010) Reconceptualizing public policy from the perspective of complexity theory. *Emergence: A Journal of Complexity Issues in Organizations and Management* 12: 52–60.
Morçöl G (2012) *A Complexity Theory for Public Policy*. New York: Routledge.
Morçöl G and Wachhaus A (2009) Network and complexity theories: A comparison and prospects for a synthesis. *Administrative Theory and Praxis* 31: 44–58.

Newman J and Head BW (2015) Categories of failure in climate change mitigation policy in Australia. *Public Policy and Administration*, pp. 1–17. Epub ahead of print 22 January 2015. DOI: 10.1177/0952076714565832.

Osborne SP (2006) The new public governance? *Public Management Review* 8: 377–387.

Ozga J (2009) Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy* 24: 149–162.

Pawson R (2006) *Evidence Based Policy: A Realist Perspective*. London: Sage.

Perrow C (2001) *Normal Accidents: Living with High Risk Technologies*. Princeton, NJ: Princeton University Press.

Pierson P (1993) When effect becomes cause: Policy feedback and political change. *World Politics* 48: 143–179.

Pilkey OH and Pilkey-Jarvis L (2007) *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*. New York: Columbia University Press.

Pressman JL and Wildavsky A (1984) *Implementation*, 3rd ed. Berkeley: University of California Press.

Rhodes RAW, 't Hart P and Noordegraaf M (2007) Being there. In: Rhodes RAW, 't Hart P and Noordegraaf M (eds) *Observing Government Elites: Up Close and Personal*. Basingstoke: Palgrave Macmillan, pp. 1–17.

Richardson GP (1986) Problems with causal-loop diagrams. *System Dynamics Review* 2: 158–170.

Richardson GP (1991) *Feedback Thought in Social Science and Systems Theory*. Waltham, MA: Pegasus Communications.

Richmond B (1993) Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review* 9: 113–133.

Sabatier P and Mazmanian D (1980) The implementation of public policy: A framework of analysis. *Policy Studies Journal* 8: 538–560.

Schuck PH (2014) *Why Government Fails So Often: And How It Can Do Better*. Princeton, NJ: Princeton University Press.

Senge PM (1990) *The Fifth Discipline: The Art and Practice of the Learning Organisation*. London: Century Business.

Sheppard JP and Chowdhury SD (2005) Riding the wrong wave: Organizational failure as a failed turnaround. *Long Range Planning* 38: 239–260.

Sieber SD (1981) *Fatal Remedies: The Ironies of Social Intervention*. New York: Plenum.

Streeck W and Thelen K (2005) *Beyond Continuity: Institutional Change in Advanced Political Economies*. New York: Oxford University Press.

Teisman GR and Klijn EH (2008) Complexity theory and public management: An introduction. *Public Management Review* 10: 287–297.

Teisman G, Van Buuren A and Gerrits L (2009) *Managing Complex Governance Systems*. New York: Taylor & Francis.

Tenner E (1997) *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*. New York: Random House.

Tilley N (2010) Realistic evaluation and disciplinary knowledge: Applications from the field of criminology. In: Vaessen J and Leeuw F (eds) *Mind the Gap: Evaluation and the Disciplines*. New Brunswick, NJ: Transaction Publishers, pp. 203–235.

Toole TM (2005) A project management causal loop diagram. In: *Paper is presented at a conference*. London: ACROM.

Van der Steen M, van Twist M, Fenger M, et al. (2013) Complex causality in improving underperforming schools: A complex adaptive systems approach. *Policy and Politics* 41: 551–567.

Van Twist M, van der Steen M, Kleiboer M, et al. (2013) Coping with very weak primary schools: Towards smart interventions in dutch education policy. OECD Education Working Papers, No. 98, OECD Publishing. pp. 2–49.

Walshe K, Harvey G, Hyde P, et al. (2004) Organizational failure and turnaround: Lessons for public services from the for-profit sector. *Public Money and Management* 24: 201–208.

Weatherley R and Lipsky M (1977) Street level bureaucrats and institutional information: Implementing special education reform. *Harvard Educational Review* 47(2): 171–197.

Weick KE (1995) *Sensemaking in Organizations*. Thousand Oaks: Sage.

Weick KE and Sutcliffe KM (2001) *Managing the Unexpected: Assuring High Performance in an Age of Complexity*. San Francisco: Josey-Bass.

Weiss RS (1994) *Learning from Strangers: The Art and Method of Qualitative Interview Studies*. New York: Free Press.

Wildavsky A (1988) *Searching for Safety*. Oxford: Oxford University Press.

Wilkins A (2014) Professionalizing school governance: The disciplinary effects of school autonomy and inspection on the changing role of school governors. *Journal of Education Policy* 30(2): 182–200.