The Brittleness of Unbundled Train Systems:
Crumbling Operational Coping Strategies

Bauke Steenhuisen¹ and Mark de Bruijne²
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, 2600 GA, The Netherlands

Abstract
This paper reports on the current status of the restructured Dutch rail sector. We empirically studied the strategies traffic controllers display to cope with daily value-conflicts in rail operations, at infrastructure manager ProRail and train operating company NS. We use a new framework to identify types of coping behavior. The findings are put in a broader perspective and related to literature on the organization of large complex socio-technical systems. In conclusion, we suggest what to make of the current coping strategies and the changed complexity of managing train systems. Current developments seem to raise the level of system performance but ignore the operational context of coping, making the train system more brittle on the long term.

Key words
Coupling, Engineering Systems, Coping Strategies, Unbundling Utility Sectors.

Abstract for the Second International Symposium on Engineering Systems
MIT, June 15-17, 2009

¹ PhD Candidate, Faculty of Technology, Policy and Management, Delft University of Technology, P.O. Box 5015, 2600 GA, The Netherlands. E-mail: b.m.steenhuisen@tudelft.nl
² Assistant Professor, Faculty of Technology, Policy and Management, Delft University of Technology, P.O. Box 5015, 2600 GA, The Netherlands. E-mail: m.l.c.debruijne@tudelft.nl
1. Introduction: Resurrection of the Dutch rail sector

Efforts to unbundle and de-regulate the rail sector are often suspected to have decreased its competences. This seemed particularly evident in the Netherlands right after unbundling as punctuality figures dipped from above 86% in 1999 to below 80% in 2001 [1]. The number of trains passing a red signal, indicating risks for derailments and collisions, increased as well by approximately 70% between 1996 and 2007, whereas the policy objective aimed to stay below the level of 1996 by 2009 [2]. A crisis manager at the Traffic Control Department of infrastructure manager ProRail considers it a miracle that the train system functions at all in the new institutional situation. The current rail operations process resembles “a muscle that first has been cut and then contracts,” he said in an interview with us.

The train system, the ‘muscle,’ is continuously run by thousands of operators, such as controllers, train staff, maintenance workers and many others. These operators ensure that trains run according to schedule. This demands tight levels of cooperation among these operators, much like muscle ‘fibers.’ Restructuring in the rail sector affected this cooperation. Train controllers that once used to work for a single company – the Dutch national railways – in one control room suddenly found themselves spread out over separate control rooms being either employed by the infrastructure manager or by one of thirty train operating companies. It seems plausible to assume – as the crisis manager at ProRail does – that unbundling the rail sector frustrated cooperation between controllers and may explain for the loss of competences.

Strikingly, however, the sector seemed to have overcome its operational vulnerability by 2007. Six years after the performance dip in 2001, punctuality figures gradually recovered and the number of passengers increased again with about 4% in 2008. These figures refer to Dutch Railways (NS), which provides more than 90% of the Dutch passenger train services. During this recovery period, NS passengers on average were increasingly satisfied year by year (NS annual reports). NS [3] even claims a worldwide top 3 position with regard to punctuality performance on highly intensive used networks. At the same time, the number of red signal passages at NS has been reduced with 20% in 2008 [4]. This favorable turn raises a paradox. How did the Dutch unbundled rail industry manage to parry its operational vulnerability?

In answering this question, we studied the daily practice of controllers in rail operations, focusing on the unbundled control rooms at the regional and local level of infrastructure manager ProRail and train operating company NS. On the basis of a large sample of open-ended interviews and ethnographic observations on site in the period of 2006 to 2008, we became acquainted with a rather brittle and ambiguously structured operational process. Despite the impression most recent performance figures provide, we argue that managing unbundled train systems has become increasingly brittle. We combine literature on ‘managing large complex socio-technical systems’ with the concept of ‘coping behavior.’

---

3 Measured train arrival punctuality with a 3 minute margin
2. Restructuring rail operations

Restructuring utility sectors has caused major concerns among critics; who point out that changes may affect the public interest(s) and important values such as safety and quality of service [5, 6]. Moreover, various international and cross-sectoral comparisons indicate that restructuring rail sectors is particularly problematic [7, 8, 9]. The underlying technological complexity of the rail sector commonly serves as an explanation for this. Since unbundling, tight technological and managerial interfaces between rails, trains, train staff and passengers are situated in inter-organizational processes. Instead of providing services via one rail organization (the old NS as monopolist), unbundling has created a new infrastructure manager, ProRail, which was split off from the former NS. ProRail allocates capacity to the NS and about thirty other train operating companies.

Perrow [10] distinguishes two important systemic characteristics in his “theory of systems and their potential for failure and recovery,” describing the manageability of these systems and the reliability of their performance. According to Perrow’s theory, technological systems without slack or buffer between their system elements require centralized planning to cope with their ‘tightly coupled’ nature. In contrast, decentralized problem-solving is required when system elements interact in an unanticipated and unpredictable manner, short for ‘non-linear interactions.’ Perrow characterizes rail technology as a tightly coupled, yet linear system. Assuming that the Dutch and the U.S. railway technology are more or less comparable in nature, and realizing that the rail technology did not change significantly in the Netherlands since the restructuring, we expect centralized planning as an apt strategy in rail operations.

Indeed, in rail operations, we find tasks highly specialized and cooperation across tasks adheres to strict protocols, signaling tight interfaces. A year ahead, detailed time schedules prescribe most train departures and arrivals accurate to the minute. Responsibilities for multiple values are distributed across many organizations that together provide these services. Controllers from different organizations jointly monitor, re-plan and smoothen traffic flows in real-time to deal with unexpected events and ensure that the original train schedule is maintained. Transport controllers at NS focus on train punctuality in four regional control centers, while other coordinators are specialized to serve either the logistics of personnel, rolling stock or passengers. In multiple local control rooms, other controllers implement the plans by coordinating the shunting trains and locomotives and further filling in the complex logistical puzzles of train operations. In parallel, ProRail accommodates a similar structure for railroad dispatchers. Controllers in four separate regional control centers maximize the use of rail capacity, making sure not to discriminate any train operating company. In seventeen local control centers, another type of controller is specifically assigned to oversee safety as well as the detailed logistic possibilities. Besides these four groups of controllers, there are many more controllers and operators, regarding maintenance, crises, information services, platform management and, of course, train management.

Communication among all these controllers has become more formalized since unbundling. For example, communication between ProRail and train operating companies is only permitted at the same level. Thus, regional controllers at NS are prohibited to phone local traffic controllers at ProRail. Next, the traffic control process is now supported by ‘silent’ communication systems.
with short, standardized messages. These messages are immediately shared among many controllers. Each controller encounters these messages at their workplace that consists of several computer screens to monitor the operational processes and the many actions of other controllers in real-time (see Figure 1).

![Local traffic controller's workplace at ProRail](photo: Van den Top)

Thus, the restructuring placed tightly coupled rail operations in an inter-organizational setting. Given the system’s underlying complexity, the current centralized planning and strict task specialization seems an appropriate management strategy. At first instance, however, this strategy led to an enormous performance dip. More planning and more protocols seemed to have facilitated a recovery. But, observing rail operations, we discovered controllers still display highly diverse behavior in response to daily conflicts, frequently opposed to their prescribed tasks and protocols.

3. Coping practices among controllers

Despite high degrees of planning, rail operations seem prone to ambiguity. It is hard to make sense of what actually structures rail operations, since controllers seem to act in highly person-, task- and situation-specific manners [11]. In the USA, too, researchers are struck by the various informal strategies by which controllers cooperate [12]. Rail operations practically never run as planned. Controllers constantly deal with multiple irregularities, disruptions and other incidents, minor and major. Consequently, intricate interdependencies between many tasks of controllers tighten up in multi-layered logistical puzzles of rails, trains, train personnel and passengers. We describe the strategies of controllers coping with many unplanned value conflicts.
Coping is understood as a response to value conflicts in daily practice. A literature review yielded a wide variety of coping strategies, including triage, creaming off [13], rituals, double talk [14], smokescreens, buck-passing [15], cycling, casuistry [16], incrementalism, hybridization [17] and many others. A framework of two dichotomies stood out in this literature revealing underlying patterns in the empirical material [18, 19]. First, a ‘multi-value’ coping strategies address a conflict, whereas ‘mono-value’ strategies protect values against conflict [20]. The second dichotomy distinguishes ‘deliberate,’ or explicit, from ‘emergent’ or implicit strategies [21, 22].

Crossing both dichotomies leads to a spectrum of coping responses in four broad types. The rational actor model perfectly aligns with one type, namely ‘deliberate multi-value’ coping. The assumption often is that such a response is, by definition, the most efficient way to create maximum outcome. Our examination of Dutch rail operations that emerges from our observations and interviews reveals a simultaneous mix of the four responses. By illustrating each of the four types of coping, this section brings in two notions. Each type of response appears to have its rationality and suitability, each within its own particular context. Second, the industries’ ability to optimize daily trade-offs in operations seems to diminish as a result of the restructured settings.

3.A. Deliberate multi-value coping

So, constant irregularities give rise to many value conflicts and require controllers to attune to each others problems and opportunities. To sketch the complexity of these conflicts, a staff manager describes that “The trade-off between two competing trains [a highly ordinary, though unplanned event] for the same track includes safety, non-discriminatory treatment, an optimal flow of trains, complying with the plan, transparency of the trade-off, predictability and the maximum utilization of capacity.” In practice, these trade-offs usually do not stand alone but simultaneously hinder or help a range of other trains to arrive, to depart or to continue driving. Moreover, the more trains involved in a disturbed situation, the more obscure the cause of the conflict generally is, as well the eventual trade-off. On top of that, controllers pay attention to extra values within these conflicts besides the values one of their manager mentioned above, example the actual (economical) interests of train companies and their customers and the ‘overall continuity’ of train traffic or system resilience. Next, safety of maintenance workers may also be part of these same trade-offs.

For example, a conductor may want to bring a group of passengers home. His train is the last leaving from station X that day. Before departing, he always waits for another train to arrive as planned. He expects a few dozens of passengers to transfer. But the other train is delayed. The delay even seems to exceed the planned margin of half an hour. The conductor phones a local traffic controller and communicates that he wants to wait some more. At the same time, a chief managing a night long maintenance shift may phone another local controller to emphasize that he

---

4 We used the ISI Web of Knowledge citation index to search for articles about coping between 1 January 1999 and 10 March 2008. The search resulted in tens of thousands of scientific articles. The majority of these articles have been published in psychological and psychiatric journals. We then narrowed our search to fields relevant to organizational behavior, as well as sociology, political science and public administration. This resulted in 500 articles on coping. A scan of these articles, including their references, revealed common features of coping across frameworks, journals and research fields.
planned for a tight schedule, having hundred men ready to work. For safety reasons, the chief says he sees no other option than to delay the morning rush when he cannot start punctually. These two requests of the conductor and the chief may conflict. An army of traffic controllers are in the middle of balancing these widely divergent values.

Immediately after restructuring, the old, informal ways to balance many interests in response to irregularities proved inapt in the newly unbundled situation, controllers explained. It appeared impossible to share the ‘warmth’ of information among each other. They used to have their ‘post-ears’ picking up the pieces for their puzzle from a constant stream of information and non-information as many different controllers used to sit in the same control room constantly thinking out loud, mixing work with chat [23]. After unbundling, information sharing became more demand driven as controllers that used to share in the same control room needed to phone other controllers to obtain this information. This, however, led to higher communication burdens, misunderstandings, uninformed decisions, parallel and conflicting improvisations, and quick-tempered, counter-productive quarrels, as controllers at both ProRail and NS recalled.

To repair information sharing processes and cooperation between control centers and to reduce communication burdens, managers with their controllers started to formalize operational responses to disturbed situations and to standardize real-time communication. An extensive repertoire of scenarios has been drawn up for the traffic control processes, prescribing what to do in case of irregularities, a delayed train or an obstructed railroad for example. These scenarios provide standard tradeoffs in terms of ‘if situation A is at hand, apply measure B,’ deliberately balancing the many values at stake beforehand.

These scenarios appeared quite successful in decreasing communication burdens and preventing disturbances to escalate, but they did not repair the old qualities of sharing information. Using standard scenarios too often makes controllers less sensitive to contexts in which conflicts are always cast. Such contexts usually consist of unique, detailed circumstances such as an extra train planned ad hoc, an unusual amount of passengers or a crowd of football fans instead of commuters. Weather conditions make a difference, too, for passengers’ comfort as well as system resilience. While controllers apply these standard scenarios as ‘trade-off proof’ solutions, controllers practically tend to refrain from inquiring into the trade-offs actually taking place. Accordingly, a line manager describes that the use of scenarios creates a “make-believe certainty, as if there is an ideal solution… [These scenarios] discourage controllers to think for themselves.”

3.B. Deliberate mono-value coping
During our study, many controllers (un)consciously reduced the complexity of their task to a single objective and perceive balancing conflicts with other controllers as an activity that should have been anticipated in their protocols & scenarios. For example, some regional traffic controllers at ProRail portrayed their work as fully compatible, single issue and with uncontested rules. “All we do is isolating delayed trains,” they say, “too late is too late.” These simple decisions rule – punctual trains get priority over delayed in case of a conflict – are considered essential to guarantee predictable and smooth processes and neutrality among train operating companies. Conflicts do not seem to exist, though. Similarly, regional transport controllers at NS focus on punctuality. These ‘commercial interests’ of transport companies are fenced off from
the safety function of local traffic controllers. As a consequence, however, NS controllers frequently cope with conflicts between delayed trains and passenger interests without a possibility to attune their solutions with the detailed logistic puzzle local controllers at ProRail deal with.

At ProRail, we observed that communication burdens between local and regional controllers were effectively reduced, but, in fact, many controllers almost seemed to work without each others expertise and information. At disturbed, busy moments, building an alternative plan along the formal communication lines could easily take 20 to 30 minutes. “This is unworkable,” controllers at NS say, “What used to be dealt with in one second now takes a whole circus of calling and calling back with the risk that you can not even reach somebody… Sometimes you loose a rail track for 10 minutes, because of consultations.” Moreover, the needs and the circumstances are simply not always known 30 minutes in advance. Therefore, communication tends to be avoided in practice. Consequently, controllers learnt to just act upon their interests, for example an NS controller may compel a train driver to wait for a train connection without deliberating with ProRail. Thus, controllers developed independent mono-value coping responses, sometimes time-efficient, sometimes catch-as-catch-can.

3.C. Emergent mono-value coping
Because many conflicts remain unrecognized among independently coping controllers, controllers structurally displace the negative consequences to other controllers instead of balancing their mutual conflicts. Their mono-value orientation implicitly generates rather undeliberate trade-offs.

The neutrality principle, for example, is directly operationalized in hard decision rules for regional traffic controllers. ‘A train is a train’ and ‘delayed trains should not hinder other trains,’ independent of the train operating company. In contrast to this non-discrimination principle, passenger interests are neither operationalized nor sanctioned for traffic controllers at ProRail. When local trains of a local train operating company competes with a delayed NS intercity train for the same rail track, traffic controllers surprisingly reverse the trade-offs they used to make before unbundling. In the former situation, the intercity train would often be prioritized over the local train, considering the passenger interests of both trains. Besides, delaying intercity trains often has much more consequences on the stability of the total network compared to delaying a local train. In the new situation, traffic controllers favor the local train over the intercity on the threat that train companies go to court. Some controllers at ProRail still try to optimize interests on a system level as they used to do, but only when operational discretion allows them to or when other train operating companies are prepared to sacrifice their train to repair a disturbance they did not cause. The many new, more commercial and more efficiently managed train operating companies tend to be less willing to adapt for the common good. Thus, safeguarding the non-discrimination principle disables controllers to cope with conflicts deliberately.

3.D. Emergent multi-value coping
There is no controller specifically responsible as a coordinator for balancing multiple values in constant conflict situations on a system level. Instead, each controller generally responds within its own action radius for its own task specifications. Interactions between all controllers may underlie a system of checks-and balances, coping with these coordination issues without
deliberately addressing conflicts. In other words, conflicts are tossed around from one controller to the other and possibly in a functional way.

In practice, many value conflicts de facto decentralize to local traffic controllers first as they have the most detailed overview of traffic flows. These local control rooms tend to be the busiest place in the traffic control process, too. Because of these conflicts, some of the more proactive local controllers take various responsibilities of other controllers on themselves to enable the traffic control process with more situational optimization, in their perception. These controllers usually find ways to balance conflicts between passenger interests, punctuality, train safety, safety of maintenance workers, neutrality, capacity, costs and various practical impossibilities simultaneously, but when work pressures peak it often happens that they lose control over the situation and do not see what is optimal anymore and whether their decisions work out positively or negatively at all. Attempts to balance multiple values in constant conflicts, thus, at the same time undermines this very strategy and risks less deliberate outcomes.

For the same reasons, to regain control over conflict situations, many other local controllers do not even attempt to balance but deliberately avoid complex coordination issues and retreat to their formal task that does not prescribe any proactive coping. This would be described as ‘emergent mono-value coping.’ At a rough estimate, confirmed by line managers, these two opposite responses are rather equally distributed among the controllers.

In sum, restructured rail operations display ambiguously structured and diverse coping strategies. Two paradoxes emerge from this coping perspective. First, we found relatively many decentralized improvisations for a tightly coupled system. Apparently, there are nonlinear interactions, too, to manage in the train system and controllers adapt to these challenges. Second, the rise of task-oriented, reactive and defensive coping strategies seems to make rail operations increasingly brittle, as value-conflicts remain unrecognized and trade-offs unmanaged. The restructured settings tend to push controllers into independent coping strategies, increasingly isolating them from value conflicts that are inherent in complex system operations. Supporting more collaborative coping responses, however, seem essential to develop and maintain situational awareness [12]. This awareness allows operators to disentangle the cause-effect chains of operations in which controllers perform only an isolated, single act. This allows controllers to recognize value-conflicts and regain their control over the many trade-offs that occur on a daily basis. Yet, this decreasing awareness, this brittleness does not show, at least not in recent performance figures. The next section further discusses these paradoxical findings.

4. Discussion: Managing complex technologies in a multi-value context

The ‘push’ towards a new set of coping strategies can be explained and understood if we understand the changed nature of managing the train system in an unbundled situation. We argue that the increasing use of undeliberate, mono-value strategies in rail operations may serve as an indicator to identify these changes.

With regard to Perrow’s theoretical framework about the management of safety in large socio-technical systems it could be argued that as a result of a range of factors, the interactive
complexity of the Dutch railway seemed to have changed significantly. At first, the increased number of organizations presented a much more complex and fragmented system that has proven far more difficult to manage and operate reliably, safely and punctually than before [24, 25]. Instead of being in charge of all railway traffic, rail operators at Prorail have to console the demands of dozens of organizations under conditions of ‘equal access’, thereby loosing much of their ability to ‘control’ railway traffic. As a result of these conditions railway controllers found themselves confronted with the need to significantly redesign many operational processes in the railway industry. Later, overall performance seemed to improve again although the operational process still appeared ambiguously structured. Some operational responses were made consciously, deviating or confirming with protocols, other adjustments to ‘cope’ with the increased intransparency and complexity of infrastructure operations were – as we have described – made less consciously.

The institutional changes doubled the number of control rooms in the sector to operate the system, thereby diffusing available information over more entities and people. This information is also less comprehensive and fragmented in bits and pieces. To nevertheless ‘seamlessly’ connect the increased amounts of organizations, their systems and their technologies, requires the use of more information technology (telephones, computer systems, etc.) to operate, coordinate and monitor the complex interactions in the railway system [26]. What emerged after the reform, thus, is a relatively more complexly interactive operational process in which more organizations with sometimes conflicting interests become involved.

The implication of these findings is that the train system has shifted towards a complexity that is less manageable. Perrow sketches a seemingly unsolvable puzzle about the way organizations should deal with complexity and coupling as it requires mutually exclusive organizational responses. We argue that as a result of unbundling railways have ‘shifted’ towards this conundrum. Interestingly, the primary driver of this shift was an institutional reform, rather than a technical change. As posited, predominantly, the technology used in railway operations has stayed the same. But still the technological system may be considered less linear – at least not from the perspective of those daily managing it. This underlines Perrow’s claim in his theory of reliability in socio-technical systems that technology and institutions must be seen as fundamentally intertwined.

The result is a managerial “catch-22”. On one hand, systems that are tightly coupled must be operated through a highly centralized authority structure, with operators reacting immediately using predetermined Standard Operating Procedures. “Decentralized systems are too slow to respond to widespread multiple failures, because the units cannot be instantly and unquestioningly controlled from the top where often there is a superior view” [27]. On the other hand, a high degree of interactive complexity of the system requires decentralized management because “[d]ecentralized units are better able to handle the continual stream of small failures” (ibid., p. 152). This research identifies and underlines these problems in a multi-value context.

- Paradox 1 is solved: due to the double complexity and the catch 22, the rail sector needs both centralized and decentralized management
- Paradox 2 seems to imply that the industry found a (temporary) solution to the catch-22 management dilemma. Solution: loose connection between plan and practice, facilitating both centralized and decentralized management.
5. Conclusion

Restructuring in the rail sector appeared to push controllers towards a more brittle operational process with a narrower repertoire of coping strategies. The change in system complexity as encountered by various traffic controllers has not been fully recognized.

So, institutional fragmentation negatively affects the manageability of large-scale infrastructures, but not necessarily overall performance. Despite new strategies of operators to deal with the increased complexity of running the railways – operational discretion, strength of diversity, fragile process with a crucial loose connection – it could be argued that this is not a stable situation, neither for managers, nor for controllers. The designed task specifications and many engineered solutions currently seem to raise the level of system performance but only seem to do so at the expense of the operational maneuverability to cope. It is exactly this crucial context of coping that enables operators in train systems to deal with unexpected events that may threaten service continuity. This context allows for decentralized operations. But the restructured settings tend to remove the more sensitive coping strategies of system operators to deal with unfamiliar and unexpected events. As the visibility of cause-and-effect chains further decreases, the chance of unexpected events increases and unseen incidents may ripple into system failure more often and more easily. In Perrow’s view, the system could perhaps be increasingly considered as a normal accident prone sector.

It remains to be seen what the long-term effects of these institutional unbundling in the railway industry are on the operational resilience of system operators and whether a ‘normal accident’ is about to happen. Relevant future research, therefore, is to keep track of the status of these operational coping strategies in the near future. Another direction is to compare these qualitative impressions with quantitative datasets on the occurrence of various, more detailed types of errors and disruptions to see whether the system stabilizes or destabilizes, like Roe and Schulman (2008) did recently for the California electrical grid.

Acknowledgment

The authors are grateful to all the respondents at NS and ProRail for contributing and the Next Generation Infrastructures Foundation (www.nginfra.nl) for funding the research.

References


