"The limits of normal accident theory"

Andrew Hopkins

Abstract

Embedded in Perrow's book Normal Accidents is a theory of normal accidents. The theory is limited in a number of important respects. First, it applies to only a very small category of accidents. Second, its concepts are ill-defined leading to serious ambiguities about just what the theory covers. Third, in some crucial respects it appears to be wrong. Fourth, recent attempts to reformulate the theory by expanding it in various ways - by incorporating basic insights from organisational sociology along with the concepts of interest group and power - actually replace rather than expand the theory. Finally, the theory is of very limited policy relevance.

1. Introduction

Charles Perrow's book, Normal Accidents, is something of a classic in organisational sociology. In what is now known as `normal accident theory' (NAT) the Yale sociologist argues that major accidents are inevitable in certain high-risk systems, such as nuclear power stations. The book has been particularly influential among researchers concerned to understand the organisational origins of disasters and the strategies which might be used to make organisations safer.

My own interest in Perrow's work stems from an attempt to understand major coal mine accidents. There have been five disastrous explosions in Australian coal mines since 1972, each claiming more than 10 lives. How is this record to be explained? Does normal accident theory provide the key to understanding the apparent inevitability of these accidents? More generally, how useful is NAT for understanding industrial and other socio-technical disasters and in providing guidance on how to make disaster-prone organisations safer? These are the questions this article seeks to address.

We must begin by asking just exactly what is NAT. Normal Accidents is a fascinating book. It is full of insights into accidents in nuclear power stations and chemical plants, collisions between ships at sea, dam collapses, mining disasters, and more. It also contains important observations on how risk can be reduced. But extracting the theory of normal accidents from this wealth of information is not straight forward. Students of law will be familiar with the distinction between the ratio decidendi, the reason for the judicial decision, sometimes not at all clearly stated, and obiter dicta, the often interesting but essentially irrelevant comments made in the course of the judgement. It is the ratio which has precedent value and the job of those who seek to make use of legal decisions is to extract the ratio from the obiter. Sifting through the obiter of Normal Accidents for the theory of normal accidents bears some similarities to this process.

……..

Perrow first defines normal or system accidents (the terms are used interchangeably) as accidents involving unanticipated interaction of multiple failures in systems with high-risk technologies (Perrow, 1984, p. 70). They are different from component failure accidents which involve the failure of one component (although this may lead to the consequential failure of other components in a predictable fashion).

In order to explain normal accidents, Perrow argues that systems can vary in two ways: they may be either linear or complex, and they may be tightly or loosely coupled.

Linear interactions are those in expected and familiar production or maintenance sequence, and those that are quite visible even if unplanned. Complex interactions are those of unfamiliar sequences, or unplanned and unexpected sequences, and either not visible or not immediately comprehensible. (Perrow, 1984, p. 78)

As for the second dimension of variability, a system is said to be tightly coupled where one thing follows rapidly and almost invariably from another with little opportunity for human intervention. Tightly coupled systems are normally highly automated systems. A system is loosely coupled when things develop more slowly, where different outcomes are possible and there is plenty of time and opportunity for human intervention to deal with problems before they become serious.

Perrow argues that where systems are both complex and tightly coupled, accidents are inevitable. Where the system involves a high-risk technology, disaster is inevitable. These are the bare bones of NAT.

2. The limited relevance of NAT

Already a major limitation of the theory is apparent: it is not an explanation for disastrous accidents in general but only those which occur in complex, tightly coupled systems. It turns out that many of the disasters which have occurred in recent decades are not explicable in these terms. The gas leak from a chemical plant which killed thousands at Bhopal in India, the fiery destruction of the space shuttle, Challenger, the Soviet nuclear reactor accident at Chernobyl from which people are still dying, the Exxon Valdez tanker oil spill in Alaska, an undoubted environmental disaster if not a human one - none of these is a normal or system accident, according to Perrow (1994, p. 218). They are no more than `component failure accidents' which cannot be analysed in system terms. ``They are alarmingly banal examples of organisational elites not trying very hard'', he says (Perrow, 1994, p. 28). In other words, NAT has nothing to say about many of the most publicised disasters of our time.

Perhaps a little surprisingly, Normal Accidents contains extensive discussions of accidents which the author argues are not normal accidents. The longest chapter in the book is about marine accidents, mainly collisions, but only 5-10% of these are system accidents he says (Perrow, 1984, p. 175). The rest are simple component failure accidents. The systems involved are relatively loosely coupled and there is a lengthy period prior to most collisions when appropriate human intervention could have averted the accident.

Likewise, in a chapter on dam, quake, mine and lake disasters, he argues that most of the systems concerned are either linear or loosely coupled and hence that these are mostly not system accidents. My own work on a recent Australian mining disaster shows that the system was indeed loosely coupled: signs that trouble was brewing were present long before the explosion, and had management attention been focussed on these signs at any stage the disaster need not have occurred (Hopkins, 1999). Perrow does, however, offer some obiter about mine disasters. ``Mining is inherently difficult and dangerous'' (Perrow, 1984, p. 245), he says, and ``it would appear that there is an irreducible hazard in mining - an unpredictable environment for humans'' (Perrow, 1984, p. 251). Put another way, mine disasters are ultimately not susceptible to sociological explanation. Here, then, is the answer to the specific question I started with: neither NAT, nor even the book Normal Accidents provides much help in understanding the recent series of mine explosions in Australia.

3. The absence of criteria for measuring complexity and coupling

Given that NAT provides an explanation for a very limited class of accidents - those occurring in complex, tightly coupled systems - it is important to be able to specify which systems are of this type. Perrow seems to be clear in his own mind about this and is able to locate a large array of systems in the four-cell space defined by the dichotomies linear/complex and tight/loose (Perrow, 1984, p. 97). But as one book reviewer pointed out, ``the absence of clear criteria for measuring complexity and coupling makes his examples seem anecdotal, inconsistent and subjective'' (Kates, 1986). Another put it more ironically: ``his constructs are loosely coupled to his illustrations'' (Roberts, 1989, p. 286).

Perrow's own discussion demonstrates these inconsistencies. Space missions he positions as complex, tightly coupled systems, yet the Challenger disaster was, he says, a simple component failure, not a system accident. The US military early warning system is positioned as a complex, tightly coupled system. Yet in the text we find the intriguing comment that ``the early warning system appears to be moderately complex and coupled, but not disastrously so.'' (Perrow, 1984, p. 291). It is hard to avoid the suggestion of circularity in this last comment. There have in fact been no disasters resulting from failures in the early warning system. It is not legitimate to conclude from this that the system is only moderately complex and coupled, as Perrow appears to do: complexity and coupling must be defined independently of the phenomenon they are designed to explain. Perrow is aware of this problem (Perrow, 1984, p. 97, footnote) but dismisses it as unavoidable.

The absence of clear criteria for measuring complexity and coupling is particularly problematic when it comes to evaluating the theory. Scholars writing about so-called high-reliability organisations (HROs) argue that these systems operate far more reliably than Perrow's theory might suggest. ``Working in practice but not in theory'' was the provocative title of one account of their research (LaPorte and Consolini, 1991). Among the systems studied are flight operations on nuclear aircraft carriers, civil air traffic control, and electricity power grids. These are all complex, tightly coupled systems, according to high-reliability theorists (LaPorte and Rochlin, 1994, p. 22). Perrow disagrees.

> Air traffic control has been able to reduce tight coupling and is basically a linear system rather than a complexly interactive one. It is not all that clear to me how coupled and complexly interactive power grids are. I judge them to be quite tightly coupled but moderately linear. . . Flight operations on the aircraft carriers are basically linear, and since landings and takeoffs can be easily stopped or delayed, the system may be seen as loosely coupled. (Perrow, 1994, p. 216)

High-reliability theorists do not claim that their work constitutes a test of normal accident theory but they do claim to be adding to our knowledge of how the chance of accidents in complex, tightly coupled systems can be reduced. Perrow's response is that their work is essentially irrelevant to his because they are studying different types of system. This dispute with high-reliability theorists highlights the absence of any criteria by which decisions can be made about complexity and coupling and thus the difficulty of engaging with NAT, let alone testing it. But more than this, Perrow's response again limits the utility of his theory by restricting even further the number of systems to which it applies. It seems that any attempt to engage empirically with NAT runs the risk that Perrow will judge the empirical situation one to which the theory does not apply.

Moreover, both sides can play at this game. Perrow (1994) is approving of Sagan's (1993) attempt to demonstrate the applicability of NAT to the US nuclear defence system which on several occasions during the Cold War years came close to launching a nuclear strike or exploding nuclear weapons by accident. LaPorte and Rochlin (1994) argue, however, that this system was linear and not complex and thus that NAT provides no explanation for these close calls (LaPorte and Rochlin, 1994, p. 223).

NAT was initially developed to explain the near disaster at the Three Mile Island nuclear power plant. Perrow argued in his book that the theory could also be extended to cover certain accidents in the petrochemicals industry. But it seems that the moment other researchers seek to apply NAT beyond its original context, doubt can be raised as to its relevance.

4. NAT and structures of authority

Let us return to the problem of specifying normal accident theory. At a late stage in Normal Accidents, Perrow introduces an important discussion about the type of authority structure best suited for dealing with crises with the potential to cause major accidents (not just normal accidents). The account uses his four-fold classification of systems based on coupling and complexity and it provides a more detailed explanation of why accidents are inevitable in complex, tightly coupled systems (Table 1). This account can reasonably be taken as an integral part of the theory of normal accidents. The argument is as follows.

Where systems are tightly, as opposed to loosely, coupled and there is little or no time for reflection on the job, authority must be highly centralised with operatives doing what they are supposed to do in a pre-determined and unquestioning manner. Provided there are no other considerations, this will maximise the prospect of disaster avoidance. Where systems are complex, as opposed to linear, central decision makers will not be in the best position to comprehend what is going on and local

……

decision makers may be better placed to avoid system failure, provided there are no other considerations.

Where a system is both tightly coupled and complex these imperatives pull in opposite directions: tight coupling requires centralised authority while complexity requires autonomous decision making. This conflict makes it impossible to devise an authority style which will reduce the risk of accidents in complex, tightly coupled systems. This explains just why it is that accidents are inevitable in complex, tightly coupled systems.

There is something logically appealing about this formulation. The critics argue, however, that it is empirically incorrect - that it is possible for authority to be simultaneously centralised and decentralised. This is, in fact, what the students of HROs have observed. According to LaPorte and Consolini (1991, pp. 31±32) HROs have ``nested authority structures''. This means that, while they routinely function in a highly bureaucratic way with rules and standard operating procedures designed to maximise operational predictability, at times of greatest pressure, authority shifts downwards to frontline operators who are accorded considerable decision-making discretion. The rationale for this is that, at such times, senior managers, who are removed from the arena of operations, will not be able to respond as quickly as frontline operators; devolving authority at critical times increases the likelihood that decisions will be taken that avoid disasters.

The difficulty about whether HROs are complex and tightly coupled does not affect this argument. HRO theorists claim that it is empirically possible to have both centralised and decentralised decision making. If so, there is no reason why complex, tightly coupled organisations could not in principle adopt such a decision-making structure. It follows that there is no theoretical reason why accidents in complex, tightly coupled systems should be inevitable.

In short, if Perrow's argument about authority types is taken to be an integral part of NAT then the evidence suggests that the theory may well be wrong in claiming that accidents are inevitable in complex, tightly coupled systems. Of course, if the argument about authority types is not integral to NAT, but more in the nature of obiter, then the core theory is unaffected by this critique.

5. Sagan's formulation of NAT

The preceding discussion underlines the importance of specifying the theory as clearly as possible. Sagan (1993) attempts to do this for the purposes of applying NAT in his study of the US nuclear deterrence system. He identifies the following propositions as the ``causal mechanisms'' (Sagan, 1993, p. 48) of normal accident theory (Sagan, 1993, p. 46):

1. accidents are inevitable in complex and tightly coupled systems;
2. redundancy often causes accidents: it increases interactive complexity and opaqueness and encourages risk taking;
3. organisational contradiction: decentralisation is needed for complexity, but centralisation is needed for tightly coupled systems;
4. safety is one of a number of competing objectives;
5. denial of responsibility, faulty reporting, and reconstruction of history cripples learning efforts;
6. a military model of intense discipline, socialisation, and isolation is incompatible with democratic values; and
7. organisations cannot train for unimagined, highly dangerous, or politically unpalatable operations.

Perrow (1994) has no difficulty with this characterisation of his work. But from the present perspective this is a highly problematic summary. Propositions 1, 2 and arguably 3 are integral to NAT. The remaining four are certainly propositions which can be found in Normal Accidents but they are not elements of the theory; they are obiter, scattered along the way. Propositions 4 and 5 apply to every disaster (Turner, 1978); they are not features of system accidents alone, nor do they have anything to do with complexity or coupling. Propositions 6 and 7 are totally remote from the key propositions of NAT. Putting this another way, testing propositions 4±7 as Sagan does has no bearing on the validity of NAT. Finding support for them in the context of the nuclear deterrent system does not constitute support for NAT and simply confuses the issue.

6. Garbage Can Theory

Perrow himself returns to the question of just what constitutes NAT in his 1994 article. He appropriates so-called Garbage Can Theory, ``to more sharply conceptualise NAT'' (Perrow, 1994, p. 216) than he did in his book. Garbage Can Theory (Cohen et al., 1988), infelicitously and inappropriately named, is the idea that organisations inevitably behave in unpredictable ways. There are three important propositions involved.

> First, . . . the organisation operates on the basis of a variety of inconsistent and ill-defined preferences. Different individuals at different levels of the organisation may hold conflicting goals; . . . organisations may not even know their preferences until after choices are made. Second, . . . although the organisation manages to survive and even produce, its own processes are not understood by its members. . . Third, there is extremely ¯uid participation in the organisation's decision making process. Participants come and go; some pay attention, while others do not; key meetings may be dominated by biased, uninformed or even uninterested personnel. (Sagan, 1993, p. 29, Sagan's attributions omitted)

All this is very familiar to students of organisational behaviour. This is the theory which Perrow says allows him to more sharply conceptualise NAT.

But Garbage Can Theory does not refine NAT; it replaces it. Garbage Can Theory predicts that any organisation operating a high-risk technology will inevitably experience disaster at some stage. It is not tight coupling or complexity which makes disaster inevitable but far more mundane processes of organisational failure. Garbage Can Theory predicts disasters will occur in all high-risk technologies, whether or not they are tightly coupled and complex. It thus gives us a handle on the disasters of our time in a way that NAT fails to do. In so far as Perrow relies on Garbage Can Theory he is abandoning NAT, not augmenting it.

7. Group interests and power

At another point in his 1994 article Perrow speaks of an expanded version of NAT (Perrow, 1994, p. 217) which incorporates questions of group interest and power. He notes that ``group interests and power pervaded my (1984) book. . . the issue, I argued, was not risk but power; the power of elites to impose risk on the many for the beneft of the few'' (Perrow, 1994, p. 217). These points were undoubtedly made in Normal Accidents, but they were obiter, not integral to NAT as it was then expounded. Now Perrow seeks to expand NAT to encompass these issues. Why would system elites not put safety frst, he asks. Answer:

> the harm, the consequences are not evenly distributed; the latency period may be longer than any decision maker's career, few managers are punished for not putting safety first even after an accident, but will quickly be punished for not putting profits, market share or prestige first. Managers come to believe their own rhetoric about safety because information indicating otherwise is suppressed. (Perrow, 1994, p. 217)

But again, this is not an expanded version of NAT; it is a replacement. Group interests and power are key social science concepts, in no way special to NAT, and by themselves are quite sufficient to explain disasters in high-risk systems, without recourse to questions of complexity or coupling. In going back to these tried and true concepts, Perrow is not sharpening or expanding NAT; he is implicitly abandoning it as a theory of accidents.

8. Policy

The limits of NAT are again obvious when it comes to questions of policy. What are the theory's implications for disaster avoidance? In one sense the answer to this question is simple. Since normal accidents are inevitable in complex, tightly coupled systems, the only way to avoid them is to abandon the systems concerned. That is the logic of the argument and that is indeed what Perrow recommends in the case of systems with the most disastrous potential: nuclear weapons systems and nuclear power stations. A further implication of the theory is that the likelihood of disaster can be reduced by decreasing the degree of complexity or loosening the coupling of a system. However, Perrow does not have much to say about this.

Given the logic of his argument, any other policy recommendations would have to be based on considerations that go beyond NAT. Hence at the beginning of his long chapter, ``Living with high risk systems'', we find the following comment:

> the basis of these recommendations rests not only on the system accident potential for catastrophic accidents, but also on the potential for component failure accidents. (Perrow, 1984, p. 305)

Here again, therefore, we are dealing with obiter, interesting and perhaps helpful ideas, but ideas which strictly have nothing to do with NAT.

9. Concluding comments

How useful, then, is NAT? The answer, I fear, is: not very. It is not a theory of disasters in general but only of a very small, and furthermore ill-defined subset of disasters or near disasters. What is clear is that it does not apply to many of the best-known disasters. Moreover, it is a theory whose central explanatory concepts are ill-defined, making it impossible to evaluate in an unambiguous way. NAT is propounded in a book, Normal Accidents, which is full of interesting insights essentially unconnected to the theory. Recent attempts to reformulate NAT so as to incorporate some of these insights, in particular the ideas of organisational sociology and the concepts of interest group and power, in effect abandon NAT as an explanation for why things go wrong in high-risk systems. Finally, NAT, understandably, has relatively little to say about how the potential for disaster can be reduced.

In so far as NAT has been something of a blind alley in the quest to understand and prevent major accidents, it may now be time to return to the main road and move on. Fortunately there are other promising avenues to explore. Reason's theory of active and latent failures is one that has proved widely applicable and very useful in terms of accident prevention (Reason, 1990, Chap. 7; see also Reason, 1997). And Turner's theory that all major accidents involve some element of misinformation has been shown to be widely applicable and at the same time useful for the purposes of prevention (Turner, 1978; see also Turner and Pidgeon, 1997). Finally, Vaughan's (1996) account of the Challenger disaster is readily generalisable. These accounts will not be described here; they are mentioned only in order to indicate that there are alternatives to NAT, well known to accident researchers, if less so to sociologists. Ultimately, it is the basic insights of organisational sociology, as formulated for instance in Garbage Can Theory, together with the even more basic social scientifc concepts of interest group and power, which provide us with the best account of why things go wrong in organisations, high risk or otherwise. Perhaps the lasting legacy of NAT is not its contribution to our understanding of accidents but the fact that it provoked the development of high-reliability theory, which really has contributed to our understanding of how disasters can be avoided. The book, Normal Accidents, of course remains a classic from which students of disaster will benefit for years to come.