Challenges facing Resilience Engineering as a theoretical and practical project

Nick McDonald
Trinity College Dublin

OVERVIEW

Thesis for discussion: The attraction of the notion of resilience engineering has more to do with the weaknesses of current alternative theoretical models to come up with a convincing demonstration of their utility to help manage the complexities of the real world, than with the real power of this theory (if it is a unified theory?) to offer that leverage. Unless we, collectively, address this problem, there may be an emerging crisis that this new synthesis promises much but delivers little.

In order to demonstrate this thesis, the argument starts with some crude and schematic comments on commonly observed weaknesses of current and recent theoretical positions (including resilience engineering). A simple (simplistic?) evaluation matrix is proposed, through which to assess the ability of such theoretical positions to demonstrate in practice that they can support the better design, management or improvement of operational systems. While it seems, not surprisingly, that theories which are either narrower in focus or less powerful tend to have been more fully implemented, it may be possible to overcome this apparent trade-off by developing theories which are more adequately grounded in a systemic model of the operation concerned. To change a system in a planned manner it is necessary to understand how that system works – this requires a system model. Following this logic, the HILAS project is an attempt, within the aviation system, to resolve the apparent contradictions between the depth and breadth of a theoretical model, its power to support change and innovation processes, and its demonstable applicability and use. Its industry lifecycle concept of innovation and change is based on mobilising a critical industrial mass around a common methodological framework to support an action research process to enable a longitudinal programmme of implementation and evaluation. It is too early to say whether this will be successful, but at the very least it, arguably, poses some of the problems which the next generation of theories within the domain of organisational resilience need to address.

STATE OF THE ART

Common pervasive weaknesses of current theories

Many theories of safety, risk, reliability or resilience (etc.) suffer from a number of

common problems that prevent the productive drawing of inferences to guide action to improve the system. These include the following:

Reliance on post-hoc analysis of past events.

While drawing of the analysis of past events in order to understand how to prevent future disasters of accidents is a necessary and legitimate process, it should not be the sole source of evidence relied upon. There are logical flaws in arguing from what has happened to what will happen. Understanding the past involves some conceptual framework or model, but how good is this framework as a guide for analysing present operations or future possibilities? Too much safety analysis draws its evidence almost solely from the analysis of past events without a properly developed system model of how the system normally functions. The corollary of this is that there are very few good longitudinal studies evaluating planned implementation and change.

Loose theoretical concepts

Some fundamental concepts of safety (etc.) science are conceptually very loose, for example:

- Inability to resolve contradictory representations of similar action sequences (e.g. is intentionally not following a procedure an example of a violation or productive sense making, or both?);
- Inconsistent definitions across the conceptual space (e.g. is error defined as intentional failure or system failure?);
- Using basic theoretical terms that are not value neutral (e.g. error, violation);
- System constructs which are defined in terms of potential outcomes without a full account of their functional role in the normally operating system (e.g. safety margins, latent conditions, barriers?).

Inability to account for the system in question

There are few good socio-technical models of operational systems. Rather, many theories rely on either:

- Generalised qualitative assessment of organisational systems (including over-extension of the notion of culture beyond its explanatory usefulness)
- Localised models of operators and technology

Wrong diagnosis or predictions

When the diagnosis or prediction supported by a theory is wrong, then, while it is not necessarily the case that the entire theory is wrong, it is the case that important fundamentals of the theory must be wrong. We have a case study of the application of TEM – LOSA (threat and error management – line operations safety audit) where the diagnosis supported by this methodology was almost diametrically contrary to the real

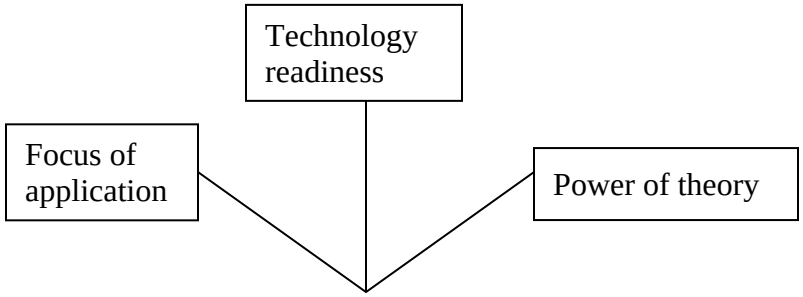
situation. The reason for this derives from fundamental definitions of error. It is not often that there is the opportunity to test the robustness of theories in this way, because the level of application and evaluation of theories is so poorly developed.

Relying on generalised metaphors as explanatory principles

Resilience Engineering has a tendency to invoke broad principles from other scientific domains as metaphors to explain underlying organisational processes – for example resonance and emergence. How do such metaphors map onto real organisational processes? For example, in relation to 'resonance', what rhythmic organisational process is being amplified by what forcing factor to create what organisational consequences? Emergence has been defined in terms of the appearance of new characteristics and qualities at complex levels of organisation that cannot be predicted solely from the study of less complex levels. In organizational terms this might be recast as the question: How do explanations of the behaviour of individuals and small groups relate to the behaviours of organisations and large socio-technical systems? Resolution of both of these issues would seem to presuppose a much tighter model of the organizational and operational system than is currently proposed. In the absence of such a model, such metaphors point to complexity and indeterminacy in such systems but do not really help us to know how to manage such factors.

An evaluation matrix

How 'fit for purpose' are current theories and models of organisational safety, resilience or risk? It is possible to draw up a matrix of three roughly orthogonal dimensions to evaluate such theories.



The **Focus of Application** dimension describes the level of the system addressed by the theory and has the following levels in increasing order of comprehensiveness:

- Operator level (individual or work group)
- Operational system – one level of activity
- Operational system – hierarchically organised levels of activity
- 'System of systems' – interactions between different major subsystems of a large operational system (e.g. civil air transport)

The **Power of Theory** dimension describes the extent to which the theoretical model enables prediction and control over the system described. This dimension has the following levels:

- Qualitative evaluation
- Quantitative evaluation
- Prediction of system performance
- Control or transformation of system performance

The **Technology Readiness** dimension describes the extent to which the theoretical model has been tested and implemented, and has the following levels:

- Theoretical concept
- Experimental / simulation / field assessment of concept
- Pilot evaluation or trial implementation of system prototype
- System fully implemented and evaluated

Clearly, in this framework, there is a desirable evolution towards a model that can demonstrate, through a full 'real-world' evaluation, its power to address complex inter-system interactions (as well as internal system complexities) in a way which enables control and improvement of the overall system performance.

COMPARATIVE ASSESSMENT OF THEORIES

Table 1, below, summarises a crude comparison between some leading theoretical models concerning organisational safety risk or reliability, using the three indices outlined earlier. It might seem that there is a trade-off between the power or focus of application of the theory and its technology readiness – those which have been applied seem to be mostly either narrower in focus or weaker in power. Indeed according to this analysis, one of the main problems centres around the Power of Theory dimension. Few theories can convincingly support a risk management process through all the stages from data collection through analysis, diagnosis, recommendations, decision, implementation and evaluation. It would be misleading to think that there is a fundamental incompatibility between these three criteria. The key to resolving this apparent contradiction is through having a comprehensive model of the operational system, which has to power to drive a cogent and valid agenda for change and innovation, and which, for this very reason, attracts the required organisational support to drive its implementation as a strategically important methodology both for the operational organisation and for the developers of new system and technologies.

The importance of a model

The fundamental key to both understanding how a system works and to being able to devise interventions to improve its functionality is to have a model of the system which incorporates the functional / causal relationships between the elements of the system.

For an industrial or operational system this is a socio-technical system – people and technology in functional unity. The organisational system (including the operational process) is as much part of the causal nexus as the technology.

What needs to be modelled is the **real system** – the objective dependencies and constraints operating in the real world independently of any one actor's intentions. Thus what is modelled is not simply a mental construction of the actors in the system but the actual material and social structure within which action takes place.

It is assumed that, arising from the necessary functional relationships which support the transformation of process inputs to outputs, that there is a generic causal structure to an operational system which transcends any particular instantiation of that system. Of course there are also an indefinite range of local variations of that generic structure according to variations on the social system, the environment and the available technology.

It is essential to be able to model the dynamic relationships between human action and system constraints. Consider the following:

- Formal procedures do not always match the real constraints in the system
- People often understand implicitly the real constraints of the system better than they do the official documentation
- People can misunderstand both the formal requirements and the real constraints of the system
- People actions can be in conformity with the formal requirements of the system and the real constraints, or such actions can be outside the boundaries of either or both.

It is important to be able to disentangle these relationships. Thus, an independent system-derived criterion of adequacy of any rule or procedure is needed. It is against that criterion that one can also judge the adequacy of any understanding of the operational situation or the appropriateness of an action.

A MODEL ENABLES DYNAMIC RISK MANAGEMENT

Having a model of the operational system then enables a comprehensive range of functions in managing and improving that system. It enables the gathering of evidence about the relationship between states of the system and operational outcomes. This is the precursor of a quantitative assessment of risk.

It provides a framework for deriving requirements for change to improve the system.

It also provides a framework to support the planning, implementation and evaluation of control measures

Table 1. Comparative assessment of theories