

Error Tolerant Communication Systems

John Aitken

Director, Aitken & Partners
Consulting Engineers

RULE BASED SYSTEMS

Railway safeworking systems are based on rules. These rules may incorporate some provision for error: for example the maintaining a headway between trains so that passing a signal at danger may not constitute an immediate risk of collision.

If the rules are obeyed and the train is run to timetable then there is no need for communication. Indeed, most railways have operated on the basis that communication is made only in exceptional circumstances and signal post telephones were invented in response to that assumption. Track circuits have been employed to control automatic signals but their operation and status was considered irrelevant to the signaller in some administrations. In earlier designs the signal post telephone was wired to a specific indicator in the signal box. The signaller knew with confidence that the driver of the train had not only stopped the train in order to use the telephone but had stopped it at a particular location. (The primary identification was the verbal statement by the driver of the location. The technology provided confirmation of the driver's assertion.) Rules were written on the basis that this technology would continue to be used: when changes were made in technology and communication systems, the rules did not follow the changes. The communications system, which had previously provided a means of identifying errors in reported position, could now become a source of error and confusion. The rest, as they say, is history¹.

Where subtle changes

The systemic approach to railway communications has gradually withered away. One might see the roots of the change in technological developments. Early last century when the majority of the higher density track was being developed, particularly in Sydney and Melbourne, the trackside communications technology used copper wires, telephone instruments, electromechanical indicators and simple switching systems, if any. (Of course, there were more sophisticated long distance systems and some large and complex exchanges but these were devoted to administration rather than train control.)

The signals and communications engineers managed the trackside systems, using the same style of technology for both signals and communications. With its strong emphasis on tradition, proven design and a resistance to change, the signalling technology was slow to embrace change. However, communications technology was in a state of rapid change, initially through miniature valves and then the transistor.

Miniaturisation and microprocessors brought changes throughout the communication industry that flowed relentlessly into the railway environment. There were concerted efforts to stop the flow with railway operated workshops creating replicas of the earlier technology: increasingly no organisation outside the railways supplied hand wound ring generators and local battery telephones. Of course, the cost of the bespoke device eventually became so great that the new communications technology was adopted and the separation of the "signals & communications" discipline into two disparate groups was evident.

... are almost unnoticed

What has this to do with rail safety? A great deal. The new breed of communications engineer looked with disdain at the large, clunky designs that originated with signals. Saw no virtue in bats eye indicators and lever switches that occupied large panels. Preferred to have multiple functions on a standard, low cost device. And, in the process, lost track of human engineering principles that had been applied and refined over the years. The problem was not evident for many years and, in the tradition of railways, was revealed only through disasters. (Perhaps those signal engineers had learned something from their mistakes.)

... in the Signal Box

In signal boxes and train control rooms the technologies sat side by side. Polished wood and Bakelite panels looked old fashioned beside the new telephone with its push buttons and display screens. As new functions were required the push buttons multiplied and gained additional functions. Code sequences and function keys generated all manner of new commands and controls. If only one could remember them! And so, subtly, low cost but obtuse new devices replaced the ergonomic, intuitive old-fashioned timber consoles.

Well almost: the old consoles remained and were in fact supplemented by the new devices and, as the obtuseness of the new devices became apparent, even more were added, with dedicated functions for some devices.

The result is evident below. The signaller at this small signal



box is confronted with six handsets, a fax machine, portable radio and the original telephone panel (grey panel to the left). What began as a carefully designed panel with a set of key switches and indicators that provided a graphical indication of the situation has been progressively degraded. Which telephone should the signaller use? Which instrument is ringing? Does he need to use a press-to-talk button (two of the handsets and the portable radio require it) or can he speak without delay? If two or three ring at once, which is important?

One could argue that the "telephone boy" has relatively little to do and should be able to manage this array of devices. However, few people are able to cope with the a continuously chattering radio, multiple telephones, discussion within the signal box and their original task of recording and

reporting train movements. The hazards associated with this arrangement are discussed in the Hexham report²; a recent IRSE paper³ and the Glenbrook report¹.

... and on the locomotive.

The inconsistency and incompatibility of train radio systems in Australia has been documented many times. Orphan systems, proprietary protocols and obsolete equipment are in use for main lines with incompatible systems for urban and long distance operations. In four cities the drivers of passenger services are unable to communicate with freight train drivers on the same and adjacent track. The "solution" that has been applied to this situation is either to ignore train radio, almost completely (the case of freight trains operating in Sydney at present), or to rely on cross connection between train controllers for the metro and long distance operations (everywhere but Brisbane). Some operators have endeavoured to fit multiple train radio systems: Pacific National's AWARE system has a "chameleon" approach, with a single user interface and many different radio systems controlled through that interface. Most other operators have multiple communication devices. For example, the Freight Australia V class locomotive was fitted with five mobile radios and a satellite telephone (pictured on the right⁴). Which of the six handsets should the driver choose in an emergency? While this is the extreme example, most locomotives have at least two handsets or microphones, in addition to a cellular telephone for each driver and perhaps a satellite telephone.



A current survey by the NSW Transport Safety and Reliability Regulator lists eleven different types of radio communication device that might be fitted to locomotives operating in NSW. Of these, five could be expected to be in use for a locomotive to be in communication throughout the state. If the locomotive operates in other states, a further set of radios will be required. Despite national agreement on a common frequency band in 1987, ARTC still operates in a frequency band used only in part of their territory. This requires a separate radio and for that section of track.

RETHINKING THE APPROACH

“If an error is possible, someone will make it”

There can be no doubt that errors will occur. A whole industry has developed to study and document human error. This has brought insights into the nature of error the mechanisms through which it occurs. Many situations can give rise to error in communications, for example: inattention due to distraction, shock and confusion. Reason⁵ categorised the types of error, distinguishing between:

- *Slip/Lapse* – an unsafe action where what was performed was not what it was intended (errors in execution)
- *Mistake* – an unsafe action purposefully executed, as intended, where the intention is erroneous (errors in planning)
- *Violation* – an additional type of unsafe intended action
- *Circumvention* – an unsafe action that is a deliberate but non-malicious violation of safety rules often done for, what is assumed, a “good” reason.

Donald Norman picks up these concepts in his book about the design of everyday things⁶ where he looks at the propensity that we have to err. Often the slips and mistakes that we make are encouraged by the design of the everyday items that we deal with. Norman’s argument is that design should recognise the reality of such error and accommodate it. He suggests that designers should:

- Understand the causes of error and design to minimise those causes.
- Make it possible to reverse actions – to “undo” them – or make it harder to do what cannot be reversed.
- Make it easier to discover the errors that do occur, and make them easier to correct.
- Change the attitude towards errors. Think of an object’s user as attempting to do a task, getting there by imperfect approximations. Don’t think of the user as making errors; think of the actions as approximations of what is desired.

The designers of a recent model luxury car could have benefited from these observations. The electronic systems in that car were so involved and so cryptic in their operation that driving became almost a secondary task. At least in this instance, the market was clear in its response. Personal computer users still suffer from many design shortcuts.

Understand the causes of Error

Norman’s suggestions are sensible and might even seem obvious. However, if one were to apply them to the railway communications systems there would be vast changes. One of the causes of error is the variety of different communication devices, with different modes of operation and different levels of coverage. Each of these devices has many modes of operation, channels for particular circumstances and locations and procedures in different situations. Most, but not all systems require a “press to talk” button to be used. Some systems require press to talk for some operations on the handset but not for others on the same handset.

It is traditional to preface each new communications contract with a survey of the “user” requirements. There have been many of these over the years and there is at least one in progress at present. Do these surveys produce the required result? One of the deficiencies of these surveys is their preoccupation with the requirements of the users at each end of the system, rather than the “systems” approach of considering all aspects of the operation of the system in both normal and degraded modes. The real risk is that the perceived needs of the users will be satisfied but hazards will be built into the system in the process.

We will briefly consider three cases where the system design can minimise the causes of error.

Design to minimise the causes – 1: Find the correct train

For example: a train driver using the Metronet system in Sydney has to “log in” to the system, entering the train trip number on a car radio size display. A train driver using the Countrynet system in Sydney has to “register” onto the system by a button press sequence. A train driver using the MDC600 system in Victoria sets an ID code that is the train trip number on a display panel. A train driver using the ARTC open channel system across the Nullarbor has no registration process. Once this difference in approach could be defended on the basis that no train driver would work outside the small section of the network where that driver had sufficient “road knowledge”. Those times have changed and the risk of error has increased. Does it matter if the driver does not perform the registration process in these systems? In most cases the registration process is essential if the driver is to hear certain types of call and is used to route the driver’s calls to the appropriate

signaller or train controller. Incorrect numbers can be entered and may not be detected, particularly if the number entered is plausible (eg the number of a train that has yet to enter the section).

Of these systems, the open channel and Countrynet are perhaps the most robust. The open channel system does not screen calls or redirect them. If the driver has selected the correct radio channel (and there is another story), the train controller will hear all conversations on the channel and the driver will hear all messages from the train controller. (In some open channel systems all parties hear both sides of all conversations). The Countrynet system identifies each locomotive that has its radio equipment turned on as an "unregistered" locomotive on a train controller workstation appropriate to the position of the locomotive (determined from GPS data). The train controller sees the locomotive number and the position of the locomotive (line and track km), together with a request from the locomotive for registration. The train controller then enters the trip number for that locomotive, associating the locomotive number and trip number. The assignment is usually followed by a voice discussion in which the trip number is identified. The strength of this approach lies in the positive identification of locomotive, the positive identification of position, graphical display of that position (by line and relative location on that line) and the visual display of all this data to the train controller. The weakness of the approach is the absence of any visual feedback to the locomotive driver. (A visual display was planned when the system was designed but has never been fitted).

A robust system? Full of visual cues and information? The designers thought so. The Hexham inquiry² showed that at least some of the information presented is either not assimilated by the train controllers or is not trusted.

Design to minimise the causes – 2: Channel Selection

Most radio systems have some form of channel selection. Where there is a channel selection or base station selection to be made, either by the signaller or the train driver, there is scope for error. And error does occur. The investigation reports mentioned previously contain examples of such errors leading to loss of communication, albeit with the best of intentions by the persons concerned. There are several techniques for avoiding or removing the opportunity for errors in channel selection. The most effective is the automated approach of Countrynet where GPS position information is used to determine the appropriate channel and train controller for each call. Metronet uses transponders to set the radio channel but these are only effective as the train passes over the transponder. Trunked radio systems deal with channel selection automatically. In these systems the "channel" is actually the address of the signaller or train controller associated with the track section and trip. GSM-R (the international standard for train radio systems) deals with the channel and controller selection automatically, in a manner equivalent to that of Countrynet. (Should one think that the channel selection issue is minor, it is perhaps worth considering that there are more than 150 channel selection combinations in use on the DIRN⁷).

Design to minimise the causes – 3: Emergency Call

When all else fails, radio communication becomes very important. These situations are usually emergencies and often the requirements for emergency calls dictate the design of the train radio system. In an open channel system there is often no concept of an emergency call. All calls are open and the person speaking conveys the emergency situation. A different approach is taken in systems (including open channel) that have selective call or dialling processes. Each of these has an emergency call button that, when pressed, signals the situation to the signaller or train controller. In most cases the emergency call message will have over-riding priority on the radio system and will be indicated to the train controller unambiguously and urgently. Countrynet and GSM-R provide an immediate display of the position of the locomotive in the emergency call indication. This information can be used effectively to initiate response and to determine which other trains may be affected by the emergency. (Though sometimes these cues are missed²).

The separate, distinct, physical emergency call button is important. An emergency call message must be readily initiated, require little thought and not be susceptible to error. Drivers and controllers react to situations of stress and shock in different ways. Although there is training and there are procedures, the investigations after the event show that shock and confusion affect those involved in train operations as much as the rest of the community. Some train driver's have assured the author that, faced with an emergency situation, their first action will be to apply the emergency brakes and their second will be to get out of the cab, as quickly as possible. It is hard to argue against this approach but one wonders whether the emergency call should not be initiated by the brake application. It would be far better to have initiated the call and then cancel it by voice than to not notify the situation and be trapped for many hours before a problem is suspected.

Make it easier to discover errors

Both Victoria and NSW have examples of manually operated radio systems being transferred to screen-based devices. However, in both cases, the screen-based device simply mimics the operation of the manual system. This has not solved the problem – a radio system user with no knowledge of radio propagation, interference ratios and channel selection is still required to make decisions about these parameters. Whether the selection is made with a rotary switch, push button, mouse or touch screen; the potential for human error remains. A better engineering solution is to redesign the underlying system to remove the hazard. This may be achieved, for example, by automating the selection algorithm for a single frequency open channel system or by removing the single frequency operation. The re-engineering may appear to be costly but the consequences of human error can be more costly.

The graphical user interface has been much discussed but often ineffectively implemented. It would seem logical, even natural for the train radio call and identity information to be displayed beside the berth that the train is occupying on the signaller's mimic display. Some systems permit this but they are in the minority. Perhaps there was virtue in the signals *and* communications approach discussed at the outset.

Screen based communication systems are being adopted as locomotives are redesigned. Some locomotive purchasers see the benefits of equipment that can be reprogrammed and updated without being removed from the locomotive. The locomotive crews are finding the advantages of having radio and channel selection automated and simplified. With larger areas available for screen display of parameters, choices and selections, it is easier for the train crew to identify and correct errors in call selection or trip number.

Change the attitude towards error

The European approach to communication system design has been one of strict standardisation, based on research into human errors. Albert Bidinger presented a paper at the IRSE in London in January 2001⁸ that reviewed the technical details of the implementation of GSM-R in Germany, with a companion paper by Christian Frerichs⁹ that reviewed the technological aspects of the cab. The System Requirements Specification for EIRENE¹⁰ (with which GSM-R alone is compliant) gives details of requirements to be satisfied for driver interfaces.

A project funded by the European Commission has investigated human factors for multicultural and multilingual rail environments¹¹. A checklist for human factors in Appendix 2 of the project report identifies considerations in the design of user interfaces (inter alia). This type of information is absent from the technical specifications for the Driver Machine Interface of the European Rail Traffic Management System (ERTMS)¹² but the derived requirements are clearly defined. The overall approach, which is one of standardisation and consistency, is contained in the UIC publication that describes the resulting interface¹³. This is a particularly useful and useable document as each element of the screen based driving, signalling and communications interface is defined. It would seem appropriate to mandate this interface as the standard for all new driver interfaces on trains in Australia.

A NEW APPROACH

Train communications systems are no longer simple voice-only radios. Rail operators demand access to data, messaging, telemetry and image transmission. When the train radio system does not provide these facilities, the users will circumvent the system, using whatever technology suits their needs at the time. There is therefore no scope for a train communication system that lacks a clear path of evolution for technology migration and development. The rapid change of technology that we considered at the start of this paper is continuing relentlessly and the expectations of users are ever increasing.

I believe it is time to take a new approach to communications in Australian railways. We need to realise that the fundamental differences between Australian railway operations and those of other countries are in fact not fundamental: they are only differences. Systems that operate in other countries can be used in Australia.

European railways with vastly different terrain, climates and cultures have agreed on an international standard: a standard that has been adopted by India and China as well as more than thirty European countries. They have adopted a technology that is already internationally standardised, has a development path and is in mass production (millions of units manufactured every year). Railway specific features have been added to this technology to complement, not change, its functions. The result is a communication system that has simple operation, automatic call routing, emergency functions, data transmission, messaging and a host of other features, all in an integrated system.

Rather than prescribe a narrow set of applications, the European platform makes a wide range of features available. Some of these features are mandated under the international rail standards, others are left to the individual operators to use or ignore. The user interface for voice-only systems is not mandated at present

but the user interface for the ERTMS is clearly defined and mandatory for ERTMS systems. The development of this interface began in 1991 with input from thirty-five train control and protection experts from thirteen European railways. The results of these interviews were translated into six experimental designs showing the information the driver needs for safe, efficient and comfortable driving. The same procedure was followed for designs for data input, for which there was a focus on the applicability of the touch screen. After the experts had commented on these designs, 150 train drivers tested them in a simulator for presentation and input. Their comments were recorded, as was their performance (brake performance, reaction times, correct answers, perceptual and mental actions). With these experimental data, the comments of the experts and ergonomic knowledge, a design was suggested that combined the best elements of the design investigated. In 1996 the combination of radio and driving functions in a single display was considered. Twenty train drivers followed a similar process with simulator tests of the proposed interface. The result was improved ergonomic consistency in the design of the graphics and layout.

One of the hidden risks in bespoke systems is the small user base and limited budget available for comprehensive testing of their error tolerance. Bespoke systems have no evolution and development path. The standardised systems have a huge international market and are used on hundreds of thousands of train trips using every day. The design deficiencies, interface ambiguities and error prone functions are therefore detected more quickly. Research and development is under way in Europe and the UK to ensure that ERTMS is error tolerant and that design flaws are detected and corrected¹⁴ before an incident occurs.

Australian railways are far from having error tolerant communications. When drivers of suburban services are able to communicate urgently with the drivers of freight services that share the same track or route in Sydney, Melbourne, Adelaide and Perth, using a standard system, then we will be on the path to error tolerant communications. Australia simply cannot afford to have yet another bespoke communication system. Error tolerant communications are essential.

¹ The Honourable Peter Aloysius McInerney, Special Commission of Inquiry Into the Glenbrook Rail Accident, Final Report, April 2001 (http://www.transport.nsw.gov.au/safety_reg/rail_investigations/index.html)

² Transport NSW, Incident Investigation, Passenger Train Collision with a Derailed Coal Train at Hexham NSW, 12 July 2002 – Final Report, December 2002. (http://www.transport.nsw.gov.au/safety_reg/rail_investigations/index.html)

³ John Aitken, Australia Wide Communication System for Rail Operators, IRSE Technical Convention, Adelaide. 14 March 2003

⁴ Photograph by courtesy of Stuart Ellis.

⁵ James Reason, Human Error, Cambridge University Press, 1990.

⁶ "The Design of Everyday Things", Donald A Norman, Basic Books, New York, 1988.

⁷ DIRN – Defined Interstate Rail Network, the standard gauge track linking the mainland capital cities.

⁸ Albert Bidinger, GSM-R, the platform for mobile communication of the railways, ie status of the project, IRSE paper read in London on 17th January 2001.

⁹ Christian Frerichs, Eurocab and the Driver MMI – an introduction to the technology, IRSE Paper read in London on 10th December 2002.

¹⁰ EIRENE Project Team, European Integrated Railway Radio Enhanced Network (EIRENE), System Requirements Specification, 15th December 2000.

¹¹ Human Safe Rail in Europe, Managing the Human Factor in Multicultural and Multilingual Rail Environments – Human Factor Analysis Techniques for Cross-Border Rail Operation, RA-97-RS-2094, 28th February 2000.

¹² CENELEC European Standard prEN50XX6-6, European Rail Traffic Management System, Driver Machine Interface, March 2000.

¹³ ERTMS (ETCS/EIRENE) MMI. The Man Machine Interface of the European Train Control System and the European Radio System for Railways. UIC, Paris, October 1998.

¹⁴ UK Railway Safety Research Program, Theme Strategy 10: Safety Critical Communications. December 2002.