Session Ten:

Assuring SIF Reliability through Function Testing – How Important is it really?

Ernst Krauss, FIEAust, CPEng, MTech
Asset Integrity Specialist, Performance Improvement (CloughAMEC)

Abstract

Safeguarding function determination is governed by Standards and supporting processes. One key element is to ascertain during the Safety Integrity Level (SIL) study a number of parameters that will enable the calculation of appropriate function test intervals, based on equipment selection, configuration and function test options. How much do we know about the competency of our personnel to carry out these tests accurately and in the way they require to be carried out to assure function reliability? How do we test and is how we test influential on the results? What must we know when carrying out the SIL evaluation and validation? Further, how to provide the assurance that we are really operating and testing within the parameters of the calculations? These are but a few questions that spring to mind when dealing with Safeguarding functions and their life cycle management. The paper examines some aspects of the inputs to the function test intervals and their impact on the outcome.

Introduction

We all have the experience of determining the SIL levels of individual Safety Integrity Functions (SIF')s and their rating in a Plant. We follow the Functional Safety standards and apply a recognised assessment methodology. Now we know what the risk levels in the plant are. For some organisations it stops here. It is not too difficult anymore to translate the SIL requirements into reality of a hardware and functional system. For many it is a matter of deciding what level of redundancy is required and then get the designers on the job to deliver the hardware, PLC and software and the final configuration. What about the calculations we carry out to validate the configuration and choice of Instrumentation? Is the practicality or otherwise of function testing in accordance with SIL requirements, configuration and failure robustness or is it a blanket testing regime? We can easily test once a year and be satisfied that we are regularly testing the SIL (or safeguarding) system. I would state now that this might not be quite aligned to the concept of the AS/IEC standards and generally the risk management requirements. What outcomes do we expect from the tests we carry out?

Potential for failures in the Safeguarding systems

The Standards[1)] define in great detail the types of physical failures that may occur. These failures are all equipment related and directly measurable. They form the backbone of our robustness calculations to ensure that we meet our Probability of Failure on Demand requirements. But how about human error or 'systematic' failures? This failure type is addressed in various ways as comment but not really spared a lot of print in the Standards. Perhaps because it is too difficult to quantify what constitutes a human error scenario and the likelihood for it to occur? A recent survey identified the following:

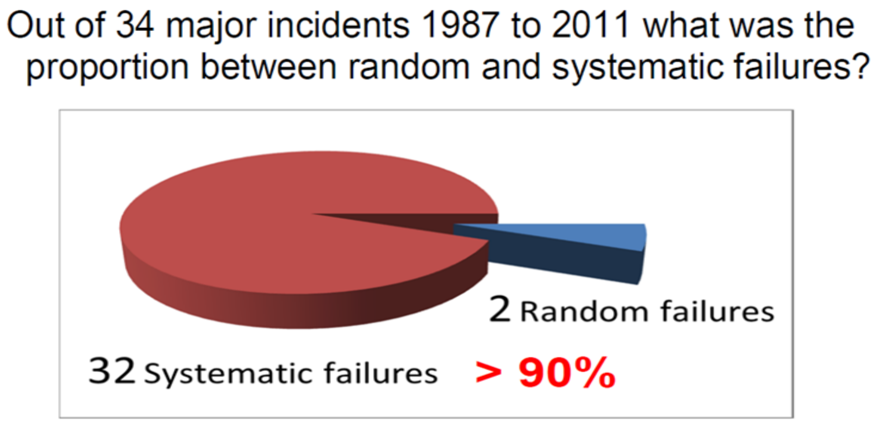


Figure 1 - System Failure distribution [2)]

If that is correct, than the majority of failures we cam expect in a SIS are systematic and procedural, competency or documentation related, not associated with hardware. This appears of great significance and in the experience of the author is a common problem in the Industry subscribing to the concept and implementing SIS. Often there is misunderstanding of elements of a SIS and how a SIF should be assessed. As the chain of events starts at the assessment stage, function testing perhaps becomes also a systematic failure fitting into the above statistic. Commonly, instrumentation loops are calibrated in preference(?) to function testing. This practice is of course not preferred for Safeguarding loops, as the individual parts or the whole loop should be function tested and not merely adjusted.

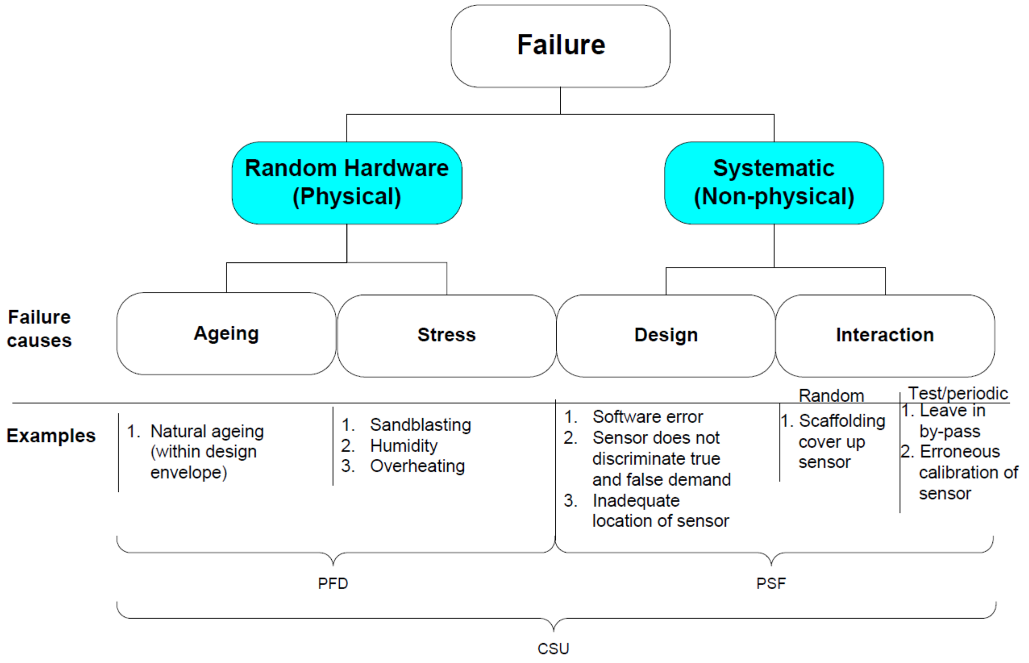

Failure
Random Hardware (Physical)
Systematic (Non-physical)
Failure causes
Ageing
Stress
Design
Interaction
Random
Test/periodic
Examples
1. Natural ageing (within design envelope)
1. Sandblasting
2. Humidity
3. Overheating
1. Software error
2. Sensor does not discriminate true and false demand
3. Inadequate location of sensor
1. Scaffolding cover up sensor
1. Leave in by-pass
2. Erroneous calibration of sensor
PFD
PSF
CSU

Figure 2 - Failure type overview [3)]

A survey in the UK (from 2002) shows that organisations have the greatest difficulty in following the spirit of the Standards to function test on line. By nature, some types of Instrumentation are very difficult to function test in situ, such as for instance Temperature and Flow. Pending on application and process, it may be impossible to raise process temperature to a point of initiating a plant trip or increase flow rates.

Over the years, various testing regimes have been adopted to meet the SIL requirements. As proof testing is difficult, often a function test is referenced and the only test implemented. A good example of this practice is the shutdown valve. These valves can normally not be fully tested as their actuation would have adverse effects on the process. And that is against the requirements of the Standards and Business in general that function or proof testing shall not create or cause additional risks and result in only minimal disturbance to normal operation. Partial stroke testing is applied to shutdown valves in certain circumstances, to prove that the valve is not frozen in situ (the normally open position).

Now arises the question of course what percentage of test coverage is achievable for instance with a partial stroke test. As the valve should not be moved further than 10% of travel because of danger of causing mechanical damage to the valve, we can't claim full diagnostic (test) coverage. AS 61508.2 Annexe A provides guidance on coverage factors. For the example, not more than 30% (optimistically) could be claimed as test coverage, even if that may cause some disagreement. The travel and reseating function in the closed position are critical, especially proper valve seating and sealing (leakage). This aspect is of great importance if such a valve is specified with a 'tight shutoff' (TSO) requirement, meaning that no leakage can be tolerated. How many organisations actually test for compliance with the "TSO" requirement when carrying out a proof test?

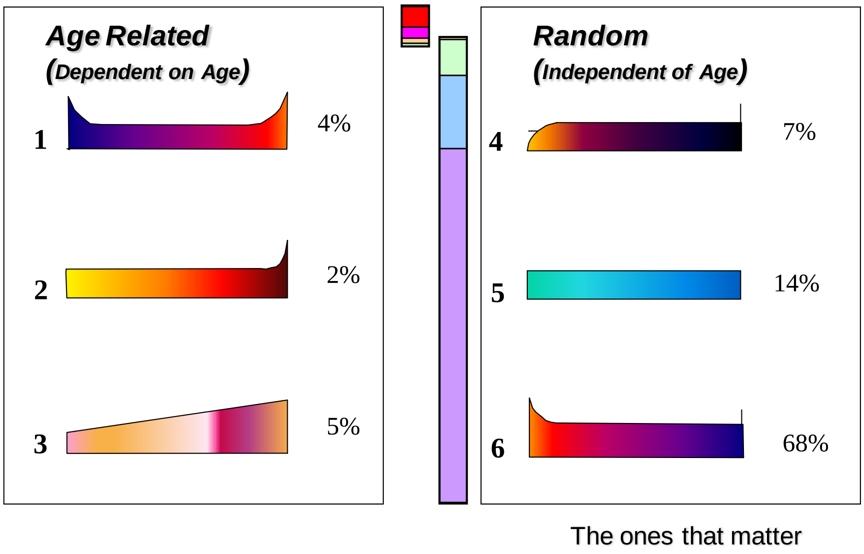


Figure 3 - Equipment Failure characteristics distribution

That leads us to consider equipment failure modes. They are the physical failures that we see in all equipment. It is generally accepted that electronic equipment is subject to random failures. As valves are mechanical devices, does the same statement hold true?

Failure mode surveys have been conducted over the years by many companies. A compilation of the results provides the conclusion that most equipment fails in a random manner [4)] , refer Fig 3.

This realisation underpins of course the need to function test (or proof test) to uncover the possible failed state of a safety system. Proof testing is therefore important as it is required to test all elements of a function. Proof testing is intended to uncover the impulse line blockage, erosion of the orifice plate, the lose connection in the solenoid valve, the seal leak in the actuator and other hidden degradation or outright failures.

Compliance Calculations and their input

The various ways to calculate the performance of the at times complex elements of a safeguarding function are explained in AS 61508.6. There are a series of possibilities mentioned. While the The simplest is the derivation of the PFD calculation methodology for low demand systems,

$$\mathrm{PFD} = 1/2 \text{ x } (\lambda\iota + \lambda_{\mathrm{PLC}} + \square_{\mathrm{FE}}) \text{ x } T$$

where:

PFD = Probability of failure on demand

λ = 1/MTBF [Mean Time Between Failure] = failure rate

λ_I = failure rate of Initiator

λ_{PLC} = failure rate of Safeguarding PLC

λ_{FE}= failure rate of Final Element

T = Test interval

As this formula is limited in its value due to not considering human errors, proof and function testing among other parameters, it should not be used for calculations for PFD compliance or function test intervals for any safeguarding functions. Fault tree methodology, Markovian models and Monte Carlo Simulation are far more suitable to determination of functional behaviour of safeguarding loops. This is explained in great detail in part 6 of AS 61508.

Calculations confirming compliance with the SIL requirements of a SIF must consider the complete safeguarding function with all its components, including the Safety PLC and software. While we are concerned about the dangerous failure portion (the unrevealed failures), we need to account for the quality of testing. This is expressed in the test coverage. Function testing a safeguarding function Initiator (e.g. a smart transmitter) may be acceptable through an electronic communicator device. How much of the input loop would we actually test in this case?

How much of a latent failure will we uncover applying this methodology? As in the example above of the shutdown valve, probably only a fraction of the potential failure modes would be uncovered. We neglect impulse lines, sensor issues and similar which are outside the capability of the communicator and its ability to interrogate the whole system. What value do we place on this test? If we strictly follow the Standard, then we would define the test coverage and describe in the SIS Life Cycle Manual the value of the function test. We also identify the measurement points and acceptance criteria for a function test. That seems like a lot of hard work for little gain. But unless we understand how this function testing contributes to functional risk management and fits into the management scope of the safeguarding functions, we can't ever expect to have confidence in carrying out the right testing and manage the impact or limitations.

We most likely can agree that a test of all elements and aspects of a SIF would be classed as a proof test; we are proving all elements of a SIF and checking its full functionality. The simplified formula does not consider the full proof test or the contribution of a function test to uncovering a latent failure. Identifying the steps required for a proof test will result in an identification of the coverage (quality and thoroughness) of the test.

To examine the impact of a testing regime, the following hypothetical SIF is used. A Markov model is used for this example as it is based on a state – space transition concept that evaluates all revealed and unrevealed ways a failure can occur and the 'transitional times' when repair is carried out. It returns a 'high precision' result, albeit at times conservative. While some people may see that as a shortfall of Markov methodology, there is a confidence factor resulting from calculations by this methodology. Of course a Monte Carlo simulation could also be used.

The typical inputs for this Markovian calculation methodology are:

- revealed failure rate λ_{rev}
- unrevealed failure rate λ_{unr}
- C_{dt} Coverage factor diagnostic test
- C_{pt} Coverage factor proof test
- T_{pi} Proof test interval
- T_r Repair time
- T_p Proof test duration
- T_{Diag} Diagnostic interval
- Q Human error factor
- b Beta factor (common mode impact)

The test durations are of some significance, as the SIF will be in a state of 'bypass' at the time, possible using less stringent methods of process monitoring. Hence there is a contribution towards probability of failure.

The Human error factor is capturing the confidence we have in our personnel to adequately carry out the SIF associated work and understand how to act and react on failure including testing and repairing. Risk assessments use a percentage as high as 10% attributing failure to activities being carried out under stress, but even for a routine activity. Various factors contribute to that state of uncertainty, such as situation, training and ability and are influenced by the work location. It is a considerably difficult area to assess. One source[5)] provides information on Operator (general term) errors of omitting steps or instructions (not intentional) as high as 5 in 60, which calculates into errors at 8% of activities. A cause for concern?

The β factor refers to the common mode failures that can be expected due to use of common elements used in a SIF. A commonly used value is 5%. This of course does not consider the failure rates of equipment; it is merely as measure of % failures that may be attributed to common item failures. High reliability components (such as factory proof tested and certified Transmitters) may have a lower factor than 5%, we often find in SIL calculations a factor as low as 2%.

Function testing impact

As mentioned before, the following calculations are carried out in a Markov model that was constructed for the specific purpose of calculation robustness requirements to achieve a specific SIL level. The calculations are examples only and should not be taken as guideline or as applicable in any process opr real life situation. The calculations are demonstrating the value of the testing regimes and how a well thought out testing regime may in fact positively impact the design of a SIF. The dangerous failure intervals are assumed rates and set at 15 years for Initiator and Final Element.

While many of the parameters are assumed, they have been taken from real life examples.

“Base Case” Calculation

The input criteria are limited to the failure rates for both revealed and unrevealed failures, repair assumed to be 8 hours (time off line to time returned to service).

A proof test interval of 10 years is used as it is highly unlikely that no proof test at all would be carried out. To carry out a test at 75% expected useful life may appear reasonable to some. For this example, there is no specific guidance provided for the proof test, its purpose, the parameters to be measured are basic (span, process variable, output signal), the tester is an Instrumentation Technician without specific training for safeguarding system maintenance. Coverage for the proof test is set at 50%. Human error is considered low, set to 1% (commonly found in SIF calculations).

Parameters omitted are:

Diagnostic test and diagnostic coverage.

The results tables in the following figures contain:

- First row – Configuration: the possible SIF configuration
- Second row - SIL achieved –based on the parameters is a SIL rating achieved
- Third row - Dangerous Failure interval in years
- Fourth row – the achieved SIF Probability of Failure to Danger

Configuration	1oo1	1oo2	2oo2	1oo3	2oo3
SIL achieved	no SIL	no SIL	no SIL	no SIL	no SIL
F_D interval [y]	14	10	43	5	37
Total PFD	3.04E-01	1.41E-01	4.64E-01	1.10E-01	1.64E-01

Figure 4 - 'Base case' calculation without testing regime

As expected in a configuration with relatively low reliability, it is not possible to achieve any SIL rating. Even the use of redundancy does not get the configuration over the threshold of a SIL requirement of 1/10. The absence of a function testing regime may be contributory to this situation. The next calculation shows the impact of a function test regime.

Identify a diagnostic test

Adding a test with a high diagnostic coverage (80%) is the next addition to the calculation set. All other parameters remain the same. The results are equally disappointing, as the proof test interval is still too far out.

Configuration	1oo1	1oo2	2oo2	1oo3	2oo3
SIL achieved	no SIL	SIL1	no SIL	SIL1	SIL1
F_D interval [y]	12	7	41	5	36
Total PFD	1.30E-01	8.67E-02	1.74E-01	8.48E-02	8.74E-02

Figure 5 - Function test added

It appears that the results in fact are worse than without any function testing, as the system is out service for longer due to the function test duration occurring more frequent.

Adjusting Proof Test interval

Generally we would not agree to have 10 year proof test interval in a low reliability system, nor should we trust a high reliability system to function without a proof test for that length of time. Introducing a 1 year proof test interval with a relatively low coverage improves the Calculated probability of failure such that we are now achieving (just) a SIL rating. Whether this is an acceptable rating for a 1oo1 configuration (Initiator and Final Element) would have to be within the policy of an organisation. In the Hydrocarbon Industry, such a close fit as in Fig 6 would not usually be accepted.

Configuration	1oo1	1oo2	2oo2	1oo3	2oo3
SIL achieved	SIL1	SIL1	no SIL	SIL1	SIL1
F_D interval [y]	10	6	40	4	36
Total PFD	9.24E-02	8.38E-02	1.01E-01	8.38E-02	8.41E-02

Figure 6 - Adjusted proof test outcome

Perhaps at this stage it is appropriate to introduce the possibility of operating with an online monitoring system, such as Measurement / Validation / Comparison (MVC). Such systems can be advantageous but require also great consideration in set up, especially when a single Initiator / Final Element configuration is involved. Today's technology still appears very costly and perhaps not enabling the monitoring of all possible failures that may occur in a SIF. A blockage of an impulse line comes to mind, although algorithms can be deployed that alert to such an event, even if it is by inference. A different set of complexities is naturally introduced by introduction of such a system.

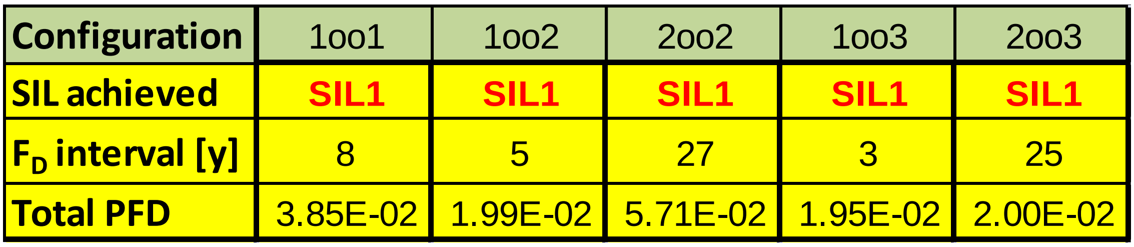
It appears that there is no 'free lunch' as the test coverage of an automated system requires even more careful consideration. Who will monitor 'the monitor'? Who is responsible for checking functionality of the algorithm and possible alarms that come up? Is the software functioning as required? These and other concerns may arise when deploying self monitoring systems.

Stepping back into the realm of 'conventional' function testing, carried out by humans, we modify the next parameter.

Adding a 6 month function test regime

Considering what many organisations might actually do in prescribing a 6 monthly function test and operating with a one year proof test, might resemble a realist approach to the SIF management question.

While the base parameters remain the same, the test coverage for both tests is revisited. Proof testing is carried out to 60% coverage, while the function test achieves only a 50% coverage. The assumptions as described above may be indicative of Industry practices.



Configuration	1oo1	1oo2	2oo2	1oo3	2oo3
SIL achieved	SIL1	SIL1	SIL1	SIL1	SIL1
F_D interval [y]	8	5	27	3	25
Total PFD	3.85E-02	1.99E-02	5.71E-02	1.95E-02	2.00E-02

Figure 7 - Introduction of a 6 month function test

Here the results confirm that the SIF comfortably conforms to a SIL 1 requirement. If we were to extend the calculation to a 1oo2 final element, we would even achieve a reasonable SIL 2 compliance. The analysis of the merits of various configurations is beyond the paper's scope, but basic conclusions on how to proceed can now be reached.

With a better understanding of test requirements and the meaning of achieved coverage there will be a further improvement of the SIF performance. We will be better enabled to manage failures.

The results bring up the question of how much coverage is needed, how much is enough? It is naturally important to consider the cost of hardware and the acceptable risk level to the organisation. Examining the results, one instinctively would say that 1oo1 configuration is sufficient. The question remains whether the proof testing can be carried out in situ or the loop elements require removing, leaving the function off line and creating a possible exposure as alternative means of fulfilling the function are required.

Depending on the general risk management philosophy, it might be decided to install a 1oo2 configuration to enable the testing of the Initiator without loss of coverage. This decision would also be influenced by the design, whether there was a provision made during design to enable an online test. Best practice would dictate that as far as reasonably practicable, online testing is the preferred testing methodology.

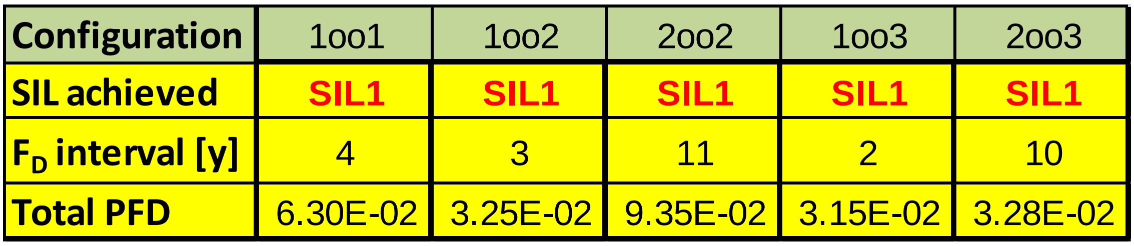
Decisions are required during the design phase of the SIF to investigate the need / option / possibility to test on line to maximise coverage and minimise risk of exposure during testing. A whole of life cost (or life cycle cost) calculation would reveal the most economical option to take. Pending this outcome, a decision can be made and must be documented in the life cycle management documentation for the Safeguarding system.

Maintenance Instructions must be adapted to clearly identify the reason for a decision for any configuration for any SIL requirements, regardless of robustness requirements. Of special importance is the documentation of required competency especially in a 1oo1 configuration. An error in proof testing or function testing could easily invalidate the functionality.

Whether there is an impact on the system performance or not due to Human error can be further investigated. The last example presented is considering a low competency, a high likelihood of Human error. For this example, the assumption is that every 6th activity will result in an error being introduced. This is akin to the previously discussed systematic error scenario, resulting in a 15% probability of incorrect activity.

Human Error impact

Fig. 8 shows the results for the calculations with a 15% Human error probability.



Configuration	1oo1	1oo2	2oo2	1oo3	2oo3
SIL achieved	SIL1	SIL1	SIL1	SIL1	SIL1
F_D interval [y]	4	3	11	2	10
Total PFD	6.30E-02	3.25E-02	9.35E-02	3.15E-02	3.28E-02

Figure 8 - Human error impact on SIF robustness

Compared to Fig. 7, the reduction in function robustness is approximately 50%! If the base calculation would show a robustness of approximately 6E-2, then human error would drive the SIF out of the acceptable range of SIL 1. Human error mitigation is a serious element in the design and management of Safeguarding Systems!

Conclusion to Proof Test practice

The investigations into the significance of designing function testing and proof testing capabilities as part of SIS design identified significant influence on the optimisation of design to meet a SIL rating. Naturally reliable hardware will improve the robustness of a SIF, but enabling testing on line wherever possible will often be achieved at comparatively little extra cost during the design, when compared with the retrofit or cost of function testing and proof testing cost if not provided.

There is debate to what extent designing of testability in situ is beneficial and where it would be useful. A HSE report from 2002 describes best practice[6)]:

- The proof test of a SIS should reflect real operating conditions as accurately as possible.
- If reasonably practicable, the SIS should be initiated by manipulation of the process variable without driving the process into the demand condition. Any approach which involves driving the process into the demand state should be accompanied by risk assessment and additional controls.
- Where process variables cannot be safely or reasonably practicably be manipulated, sufficient confidence in the correct operation of sensors should be gained by other means, such as comparison with other measurements.
- The inherent difficulties associated with testing valves and other Instrumentation (such as flowmeters) should be addressed during the design phase of SIS and additional provisions such as corroborative measurements should be made where necessary.
- Proof tests should address the necessary functional safety requirements of SIS, including functions such as response time and valve leakage class

In line with the requirements of AS 61508 and 61511, effective SIS proof testing needs to confirm the correct operation of the sensing element(s) and actuating devices, including the software functions. The most satisfactory test of a system will manipulate the process variable in order to achieve a full end to end proof test. However, practicality of achieving this outcome is very much dependent on the nature of the process, the process materials and associated risk, and on the tolerable upsets to the process and to production. Considering the difficulty experienced in practice of testing the function in line with the desired state, there is a clear need to address some of the issues and barriers to good proof test practice during the design phase.

To think that all issues can be resolved and proof and function testing can ever be fully on line and without upsets of process or causing exposure to risk during testing is illusory. What we must strive for though is the full and comprehensive understanding of how a SIF is required to operate, how to ensure competency of personnel and assure the effective management over the life cycle of a SIF.

Without designers taking on the challenge of developing test facilities for all those SIF's that reasonably can be tested online or in situ there will always be a higher degree of uncertainty about the effectiveness of SIF function in the process environment. Appropriate documentation of design features, of why function testing must be carried out in a certain way (and there are no shortcuts) and how proof testing is to be carried out, is essential for well managed SIS.

A last point on documentation. It makes sense to collect test data and use it to compare with previous test results and for performance trending. While a requirement, it is more often than not found in practice that such reports are incomplete or are not recording the required data. Recording test results, no matter at which level, is an integral part to assuring integrity of a safeguarding system.

References

1) AS 61508 / 612511 (2006 to 2011 edition)

2) I&E systems research, ref. Mirek Generowicz

3) Reliability Prediction Method for Safety Instrumented Systems, SINTEF 2003, STF38 A02420, P. Hokstad and K. Corneliussen

4) Asset Management Excellence, JD Campbell A Jardine and J McGlynn, CRC Press T5ayl;or and Francis, 2011, ISBN 978-0-8493-0300-5

5) Probabilistic Risk Assessment of Engineering Systems, M.G.Stewart and R.E.Melchers, First Edition 1997, Chapman and Hall

6) Principles for proof testing of safety instrumented systems in the chemical industry, HSE UK report 428/2002