What is Technical Integrity and How to Measure It?

Michael Edwards
Defence and Systems Institute, University of South Australia
edwmt002@mymail.unisa.edu.au

Copyright © 2011 by Michael Edwards. Published and used by INCOSE with permission.

Abstract. The impact of formal technical regulation and certification frameworks on the technical integrity of complex systems is poorly understood. In researching this problem the author is seeking a means by which to measure *technical integrity*. *Technical integrity* is defined and discussed, in both the context of complex systems for the Australian Defence Force and in wider civilian and military contexts. *Technical integrity* is defined as being composed of two separable attributes that reflect the intended, desired emergent properties of a system and the minimisation of unintended, undesired emergent properties. An approach for formulating a composite measure of *technical integrity* is developed from this definition and is presented for critical review.

Introduction

The author's area of research is concerned with complex engineered systems and inquiry into the impact of technical regulation on the technical integrity of these systems particularly during the design and development phase of the system lifecycle. A framework for this research was reported in (Edwards 2009) and this paper presents part outcomes of that ongoing research.

In the Australian Defence Force (ADF), *technical integrity* is the term used to encompass the concepts of a system's fitness for service, safety and compliance with regulations for environmental protection, i.e. its compliance with technical standards and the minimisation of unintended consequences. *Technical integrity* has various different meanings and has synonyms and related concepts occurring in other civilian and military system contexts. The first part of this paper briefly reports the results of a survey of these various meanings and related concepts in order to establish a working *technical integrity* definition that is useful across a range of complex engineered systems.

At this time, the intended positive impact of technical regulation on the technical integrity of a complex engineered system remains largely based on the application of heuristics of what should be done but without quantitative support. One of the key issues in establishing such support is the need to establish some sort of measure for *technical integrity*. With a measure established it would be possible to establish the benefit (i.e., the amount of additional *technical integrity*) obtained by the application of technical regulation, or the degree thereof. The current phase of the research is thus focused on finding support for the hypothesis: *Systems developed under Technical Regulation have higher Technical Integrity.*

The second part of this paper therefore tackles the problem of establishing a measure of technical integrity by first assessing existing methods for measuring similar properties of a complex engineered system and then by selecting a form of measure and formulating a new

approach for a composite measure of technical integrity. This approach is presented for critical review.

What Is Technical Integrity of a Complex Engineered System?

There are many definitions for *technical integrity*, many related concepts, usages of terminology and stakeholders for the desired and undesired emergent properties of complex engineered systems. Some of the related terms and concepts include: suitability for service; safety; certified products; system effectiveness; airworthiness (in the aerospace domain); systems assurance; fitness for purpose; technical assurance; product quality; compliance with class rules (in the maritime domain); compliance with build standards; mission assurance; and dependability among many others.

Some definitions include concepts of cost or economic value, while others are independent of cost or economic value. Typically all involve the concept of avoidance of unintended negative consequences (particularly safety). Typically most involve the concept of doing what was intended to be done (fitness for purpose/service, quality, mission assurance) and to have that done for most of the time (available, reliable).

Technical Integrity in the Australian Defence Force

ADF Materiel. In the Australian Defence Force, *technical integrity* is the term used to encompass the concept of a system's fitness for service, safety and compliance with regulations for environmental protection (ADF 2010). Materiel is defined as "All items including ships, tanks, self-propelled weapons, aircraft etc and related spares, repair parts and support necessary to equip, maintain and support military activities without distinction as to its application for administrative or combat purposes."

The three main components of *technical integrity* are further defined as:

- Fitness for service: The systems ability to satisfy operational requirements.
- Safety: Freedom from those conditions that can cause death, injury, or occupational illness.
- Environment Protection: Poses no hazard to the environment.

Concepts of Technical Integrity in a Wider Context

The term *technical integrity* is not commonly used across all industries or government enterprises where complex systems are engineered, maintained and put to use. This section initially looks at a few examples where the term technical integrity is used and compares these with the usage in the ADF. Then a number of related concepts and the terminology used is presented and contrasted.

Technical Integrity in the Commercial Industry. *Technical integrity* or just *integrity* is not widely used in commercial industry. One area where it is used is in the area of process plants. For example, Woodside uses the definition "Integrity – achieved when risk of failure that endangers personnel, environment or asset value is tolerable and is reduced to ALARP" (Hackett 2008) and PI Asset Integrity Ltd state "Technical Integrity in any process plant is essential to ensure that there will be no harm done to people or the environment. Assuring Technical Integrity of your facility by regular interval audits makes good business sense". (*Technical Integrity Audits* 2007). (Bale and Edwards 2000) definition is "Technical integrity is concerned with the development of the design such that it is carried out by well trained

personnel, who have been assessed to be competent, in accordance with recognized, sound practices and procedures and such that there is adequate provision by way of reviews and audits, to ensure the design intent is unimpaired in any way that could cause undue risk or harm to people or damage to the environment."

Interesting in these definitions is the direct parallel with the ADF's components of personnel safety and environmental compliance as well as the contrast between the ADF's Fitness for Service and the concepts of "asset value" and "good business sense". Clearly the latter may be explained by the mission outcomes as required by a commercial entity. That is, creation and maintenance of value to further the business.

Objectives in Regulations for Design and Operation of Systems. The design and operation of significant and complex systems is often regulated to prevent unwanted, unintended consequences by governments or by international organisations with multi-national government support. The objectives of these regulations may be thought of as technical integrity objectives. Examples include:

- Lloyds Rules Ship Classification: Compliance with a set of "rules" or other "technical standards" that have been shown to be appropriate for the objective of safe and compliant ships.
- The U.S. Department of Energy's policies and guidelines for acquisition of capital assets which include the regulations for acquisition of highly safety and security critical nuclear power assets (DoE 2006, 2008).
- The United States Navy's platform level and strike force (formally battle force) level Naval Warfare Systems Certification Policy (NAVSEA 2005) has as its stated technical objectives "Mature, Safe and Reliable Warfare Systems".

In the above few examples we see again the concepts of safety, compliance, maturity and reliability as technical objectives of regulation. A larger range of technical regulations have been examined in order to define a basis for study and to structure an ontology of technical regulation as reported in (Edwards 2010).

MIL-STD-882C Safety Compared to ADF Technical Integrity.

MIL-STD-882C (DoD 1993) defines Safety as "Freedom from those conditions that can cause death, injury, occupational illness or damage to or loss of property or damage to the environment" and identifies "operational effectiveness, time and cost" as constraints. There is close to an exact parallel when compared to the ADF's *technical integrity:* "damage to the environment" and "Freedom from those conditions that can cause death, injury, occupational illness" in MIL-STD-882C is the same respectively as Environment Protection and personnel safety as defined in the ADF's technical regulations (ADF 2010).

The one point of difference is between the "freedom from those conditions that can cause ... damage to or loss of property" of MIL-STD-882C and the "Fitness for Service" concept of the ADF's technical regulations. The latter is fully inclusive of the MIL-STD-882C concept, but as we have seen in previous sections of this paper also considers the achievement of regulatory requirements and capability requirements. Failure to meet the latter requirements in particular can be treated as a risk to mission achievement or as a mishap where the outcome is the inability to conduct certain missions or aspects of the intended mission.

A Definition of Technical Integrity

Consideration of the technical integrity and system safety concepts reviewed previously lead to a conclusion that there are two separable aspects to technical integrity of a complex engineered system. The bases for the separation are the emergent properties of the system – "useful" and "destructive" emergent properties (see (Edwards 2009), Table 1 that extends the work of (Honour 2007) to consider the influence of technical regulations on system emergent properties). The "useful" basis implies that the system meets all the requirements and standards that have been prescribed for it. We term this property of the system as the *Technical Integrity* attribute *Capability*. The "destructive" basis implies the broad notion of minimising the unspecified, but clearly undesired properties of the system. We term this property of the system as the *Technical Integrity* attribute *Dependability*. This definition is illustrated in Figure 1.

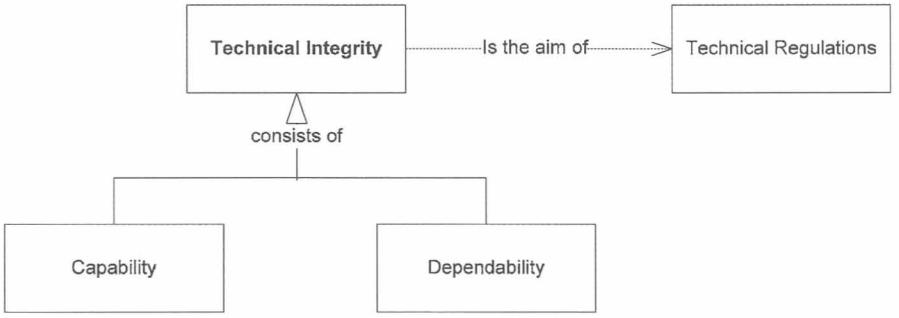


Figure 1. The Definition of Technical Integrity

More directly, *Technical Integrity* is defined as consisting of *Capability* and *Dependability*, where:

Capability is taken to mean that the system meets all specified requirements for performance, function and quality attributes in the defined environment and when used as intended. *Capability* is the designed set of emergent properties of a system that are intended and useful (i.e., required); and

Dependability is taken to mean that the risk associated with unintended, undesired emergent behaviour of the designed system is acceptably low and that undesirable limitations resulting from intentional trade-off during design are acceptable. It includes concepts of robustness (system is tolerant of usage that may exceed specified conditions without catastrophic consequences), reliability (failures do not happen often), and system safety (risk of unidentified hazards and failure modes resulting in undesirable outcomes is minimised).

It should be noted that these terms may be used differently in different contexts. *Dependability* is often used in the literature but without a universally agreed definition. For instance, (Prasad 1998) reviews many definitions of *dependability* before arriving at a definition useful for his work, albeit with some differences to that defined herein.

The definition proposed in this paper can be seen to encompass the scope of *technical integrity* as used in the ADF. Fitness for Service is equivalent to *Capability* and Safety and Environment Protection are hyponyms of *Dependability*.

Can Technical Integrity Be Measured?

Motivation to Measure Technical Integrity

Having defined *technical integrity*, the remainder of this paper is inspired (or perhaps cursed?) by Lord Kelvin who is reported to have said: "I often say that when you can measure what you are speaking about, and express it in numbers you know something about it, but when you cannot measure it, when you cannot express it in numbers your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely in your thoughts, advance to the stage of science whatever the matter may be".

The hypothesis being investigated by the author is that: *Systems developed under technical regulation have higher technical integrity*. To find support for this hypothesis, it could be imagined that a number of case studies could be evaluated for the *technical integrity* of the produced system and this could be plotted against the amount of technical regulation that was applied to the system development. Data supporting the hypothesis may look like that presented in **Erreur ! Source du renvoi introuvable.**.

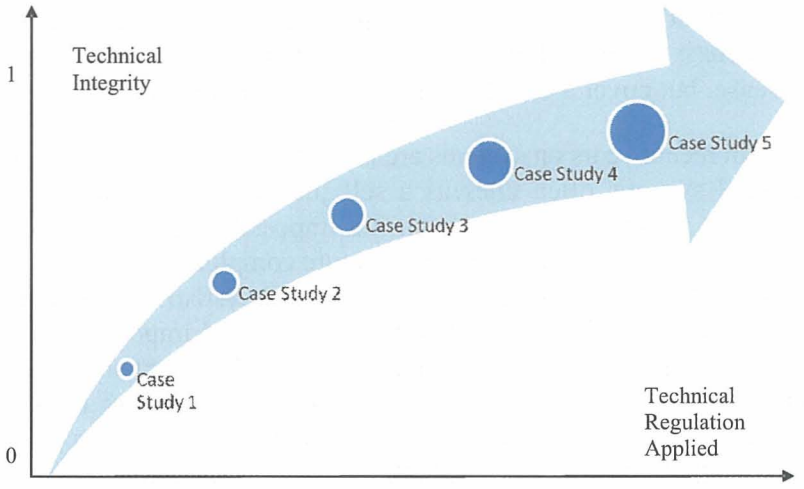


Figure 2. Concept of data supporting the hypothesis: *Systems developed under technical regulation have higher technical integrity*.

Fundamental therefore to any attempt of quantitative research into the impact of technical regulation is the ability to measure in some sense the resultant technical integrity of a complex engineered system.

Can we settle on a measure, a set of measures or some composite measure that directly indicate the *technical integrity* of the system? The measures need to be general enough so that they are applicable across the class of systems of interest and be able to be supported by quantitative or qualitative data that can be practically obtained for those systems. The measurement goals that have been set are: the measure(s) is objective; ideally the measure forms a ratio scale, but at least forms an ordinal scale; the measure is applicable across a wide range of complex engineered systems.

What Support for Such a Measure Already Exists?

Measures of System Properties. It has long been posited that there is no single overall measure for the quality or integrity of a complex engineered system (see for example (Boehm 1978)). "System Effectiveness" for a complex system, to the extent that it can be thought of as a partial proxy for *technical integrity*, most often appears as a combination of capability, reliability and availability in a value hierarchy (Parnell, Driscoll, and Henderson 2008). While, a system effectiveness measure is usually constructed for supporting selection decisions between alternate systems there is little in the way of comparative measures available to compare the utility of disparate systems.

(Hitchins 2003) nominates a set of common candidates for measuring the value of a system. His candidates include: life cycle cost; efficiency; effectiveness; performance; availability; survivability; process (e.g., simplicity, resource consumption, ...); product (e.g., utility, quality, fitness for purpose, …); entropy (the degree of disorder in the system). (Barbacci et al. 1995) also provides a useful set of software quality attributes encompassing performance, dependability, security, and safety. In "Measuring the Life-Cycle Value of Enduring Systems", (Browning and Honour 2008) take the view that a system's value can only be determined by the stakeholders who receive some utility from the system. These examples all present a view that there is a range of system qualities or attributes that should be able to be measured in some sense, but cover a broader concept than *technical integrity*.

The increasing focus in recent years on systems architecture as a technique in the development of complex engineered systems often cites as a self justification that the measurement of system attributes akin to *technical integrity* is, if not impossible, at least not tractable. For example, (Maier and Rechtin 2009) state that "architecture complements science where science is weakest – dealing with immeasurables of highly complex systems", and "most important factors are not measureable: worth, safety, affordability, political importance, environmental impact ... and have no realistic basis for numerical analysis." The problem of dealing with the attributes of "ultra-quality" systems is also salient, for instance how to measure 100% availability, zero defects, etc. It is posited in this case that measurement can only be indirect, that an analytic design approach alone does not achieve ultra-quality, and heuristic approaches must also be applied

"Qualities do not have to be expressed 'qualitatively' – they should be quantified". (Gilb 2005) dedicates a chapter of his book, about the methodology of "Planguage", to scales of measurement. He tackles concepts often claimed to be immeasurable like "user-friendly" "Environmentally-friendly" and "Customer satisfaction" by specifying scales of measure. He posits principles including "If you can't define a scale of measure, then the goal is out of control" and "If you can't quantify it, you can't control it". I would therefore speculate with some confidence that Gilb's views on technical integrity would be that "defining a scale of measure is a reasonable practical process".

Proxies for Technical Integrity. In considering the measurement of *technical integrity* it is worth considering if there are any other measures that are offered as a proxy for *technical integrity* or at least could be considered as a valid full or partial proxy. One particularly interesting set of data suggests consideration of accident rates of a complex system in service be a proxy for *technical integrity*? Figure 3 shows an example of the yearly number of accidents being used to demonstrate the positive impact of technical and operational regulation on the fatal accident rate of ADF aircraft.

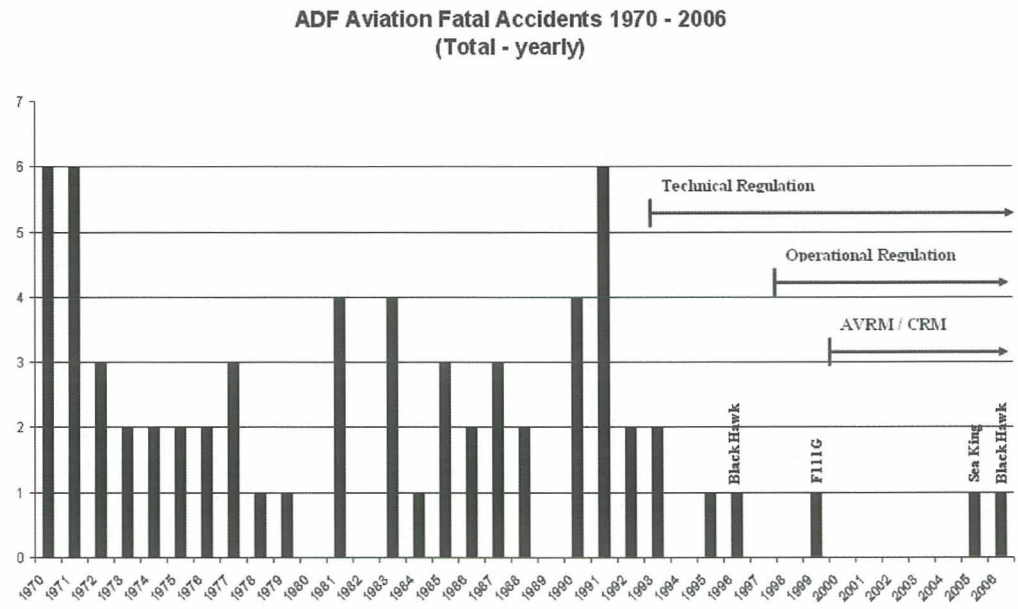


Figure 3. Total number of ADF fatal aviation accidents over the last 37 years (RAN 2007)

It is feasible that these numbers could be taken to form a measurement of the accident rate of the entire ADF aviation capability over a defined time frame and that this could have some validity as a proxy measure for *technical integrity*. However, it is worth noting that this type of measure may be easily confounded by issues like the numbers of aircraft in service in any particular year, the technology of the aircraft in service, the experience level of aircrews, the type of flying (operations versus training), sustainment spending and perhaps other variables. All these would need to be carefully controlled for in using this as a proxy measure.

Measurement of Technical Integrity is Feasible. There is established precedent for decomposition approaches to evaluate system properties and proxy measurements for properties of complex systems that are broadly similar to *technical integrity*. While it may not be straightforward, we have taken the view that measurement of *technical integrity* is feasible and desired and we therefore progress to seeking such a measure.

Selecting a Form of Measurement

This section looks at selecting a form of measurement for *technical integrity*. Measurement theory (see for example (Sydenham 1982)) has been considered in the development of the approach but the formal mathematical background and constructs are omitted for brevity in this account.

Forms of Measurement. The following forms of measurement are available for consideration:

- Direct measurement. The class of measurement including extensive measurement where something like a physical attribute of an object can be directly observed (e.g., a linear measurement).
- Indirect measurement, including:

- Derived. Establishing a relation between two or more things that can be directly measured (e.g., density which is established from the direct measurement of volume and mass).
- Associative. Measurement of at least one other property and the relationship of that to the desired property (e.g., measure resistivity to determine temperature).
- Multi-dimensional. A collection of component vectors without combination (e.g., a vector that defines the lengths of the 3 sides of a triangle).
- Conjoint. A combination of component properties which establishes an ordering of the desired property based on measurable component properties (e.g., the establishment of a "comfort" measurement by combining temperature, wind speed and relative humidity measurements).

Selection of a Measurement Form. We now consider the available forms of measure and the suitability of those for *technical integrity.* Clearly a direct extensive measurement of *technical integrity* is not possible so direct measurement is ruled out. Consideration of indirect measurement forms concludes that:

- Derived measurement is rejected since there is no known law for combining any measurable properties of a system to evaluate technical integrity.
- Associative measurement is rejected since there is no empirically known relationship between technical integrity and any other directly measureable property of a system.
- Multi-dimensional measurement is rejected since it cannot produce an ordinal, interval or ratio scale that is desired for the intended research.
- Conjoint measurement is accepted as the most appropriate measurement form for *technical integrity*. It requires some decomposition or at least identification of factors that contribute to *technical integrity* and the definition of mathematical relationships to combine these measurements into an empirical order.

Conjoint Measurement Approach

It is to be noted that we have not identified any empirically and objectively determined combination of system attributes that we can say represents *technical integrity.* But this should not be an obstruction to the formulation of a useful measure. We do however need a formalism based on measurement theory and logical arguments. The theory for conjoint measurement is not new, having been established by (Luce and Tukey 1964) and extended to include polynomial conjoint measurements by (Tversky 1967). There are no hard and fast rules for arriving at a conjoint measurement, but the usual method is to form a product structure (Roberts 1984) and assess that structure against well defined appropriateness conditions.

Product Structure. The product structure for this conjoint measurement is represented by the components of *technical integrity* previously defined, i.e., *technical integrity* consists of the system emergent properties of *capability* and *dependability*. If we represent the measures for *technical integrity* to be *TI*, *capability* to be *C* and *dependability* to be *D*, then we expect (assuming for now *capability* and *dependability* are measurable[1]) that:

$$TI = f(C, D)$$

[1] And if they are not measurable the product structure may be decomposed further.

We can think of this function as transforming any ordered pair of measurements *(c,d)* into a number on an ordinal, interval or ratio scale for TI.

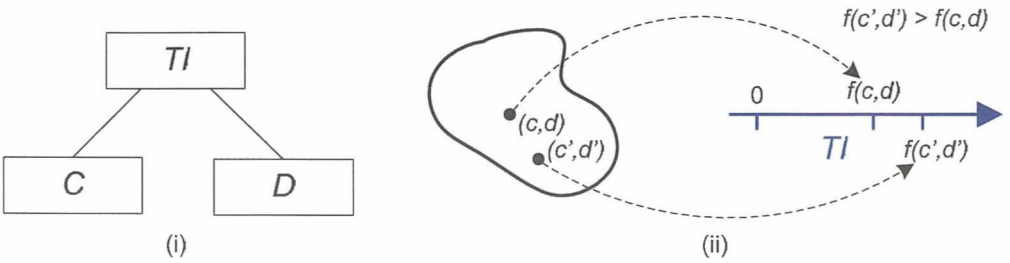


Figure 4. (i) Product Structure for TI.
(ii) Mapping of measured attributes to ordinal scale of TI.

Appropriateness Conditions. Decision theory provides advice on conditions that should be satisfied by the attributes selected for the product structure (Keeney and Raiffa 1976). They include *completeness*, *operationality*, *decomposability*, *non-redundancy* and *minimum size*. The *completeness* criterion requires consideration of the adequacy of the measurands as a set in including all the areas of concern for the intended conjoint measurement. This criterion can only be considered as a guide. The *operationality* criterion requires assurance that each measurand has a real contribution to the final measurement value. The *decomposability* criterion guides us to ensure that the measurands are independent (a necessary mathematical condition for additive conjoint). The *non-redundancy* criterion requires the elimination of double counting. The same parameter should not be measured more than once. The *minimum size* criterion guides us to limit the number of attributes chosen - more than 5 measurands is considered too many.

In order to assess the proposed product structure against these criteria we next examine what we expect the characteristics of the *TI* conjoint measurement to be, then compare this construct against the criteria.

Expected Characteristics of TI Measure. What might a measurement function for $TI = f(C,D)$ look like?

Consider a system that has met all its specified requirements. This would have a high or maximal value for *C*. However if this same system exhibited emergent behaviour that caused the death or injury of personnel, it would have a low or minimal value of *D*. What would we expect the measure of *TI* to be in this case? Clearly this would be a low value for *TI* – a system that meets its capability objectives yet often kills unintended personnel is next to useless! At the other end of the spectrum consider a system that never kills, injures nor poses any health hazard to personnel, equipment or the environment – this would have a high or maximal value of *D*. This system however never achieves its service requirements. It may be that it is just largely non-compliant with all capability requirements. This system would have minimal value on the *C* scale. Clearly this also would exhibit a low value for *TI* – a system that does not meet any performance objectives is worthless. Clearly, the desired outcome of a high *D* value and a high *C* value would exhibit a high *TI* value. These examples lead us to expect that on a graph of *D* versus *C* there would be a region of high *TI* values (desired) and a region near the axes of low *TI* values. There are then is range of intermediate values for *TI*. Figure 5 shows what such a graph could look like. The lines on the graph represent isolines of *TI*.

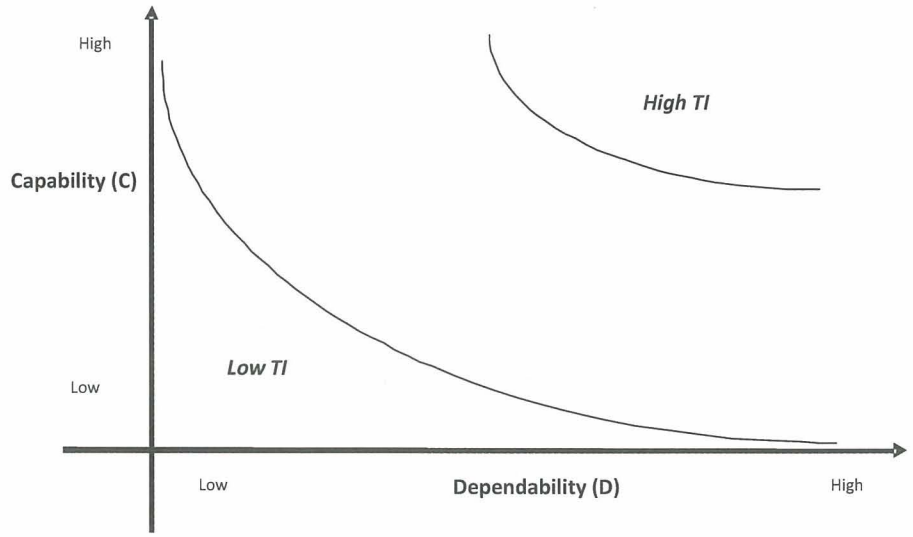


Figure 5. Expected form of *TI = f(FS,PS)*

We now assess expected measurement outcome against our completeness criteria *completeness*, *operationality*, *decomposability*, *non-redundancy* and *minimum size*. Based on work presented previously in this paper that defined *technical integrity*, we conclude that *C* and *D* form a complete set of measurands for TI. Each measurand is also clearly operational on the outcome of *TI*, e.g., a low value of either measurand is required to yield a low value of *TI*. Are the measurands independent? At this stage we assert that they are based on the intention that they separately assess intended versus unintended properties of the system. Double-counting is a potential difficulty with this approach. For instance, if a reliability requirement is well specified as Mean Time Between Failure in the system's specification then it could be double counting to consider reliability as a component of the dependability measure in this context. This will need further consideration in the definition of measurements for *C* and *D*. Minimum size has been achieved at this level of the product structure by having simply two contributing attributes. We conclude therefore that it is appropriate to form a conjoint measurement of *TI* from the attributes *C* and *D*.

Establishment of the Conjoint Measurement Model. Commonly, a conjoint measurement takes an additive form. An additive conjoint measurement for *TI* would be of the form:

$TI = \alpha.C + \beta.D + \gamma$ where α, β, γ are constants.

If we consider the boundary conditions of the measurement where we require that: if *C*=0 then *TI*=0; if *D*=0 then *TI*=0; and if *C*=maximum and *D*=maximum then *TI*=maximum, then the additive form is rejected since it will not yield these values irrespective of the values defined for α, β, γ.

The boundary conditions really suggest a multiplicative model to ensure $TI = 0$ if either $C = 0$ or $D = 0$. The mathematical foundations for multiplicative conjoint measures were first established by [Roskies 1967] as an extension of the additive conjoint model.

The following conjoint measurement model is proposed for *TI*:

$TI = C^{\alpha}.D^{\beta}$ where α, β, are constants > 0, values to be determined.

In order to demonstrate this model we constrain C, D to be real numbers in the range $[0,1]$[2] which in turn constrains TI to the range [0,1]. If we consider when α, $\beta = 1$ as a starting point we can visualise the resulting mapping of C, D to TI when it is equally influenced by its attributes. This is shown in Figure 6.

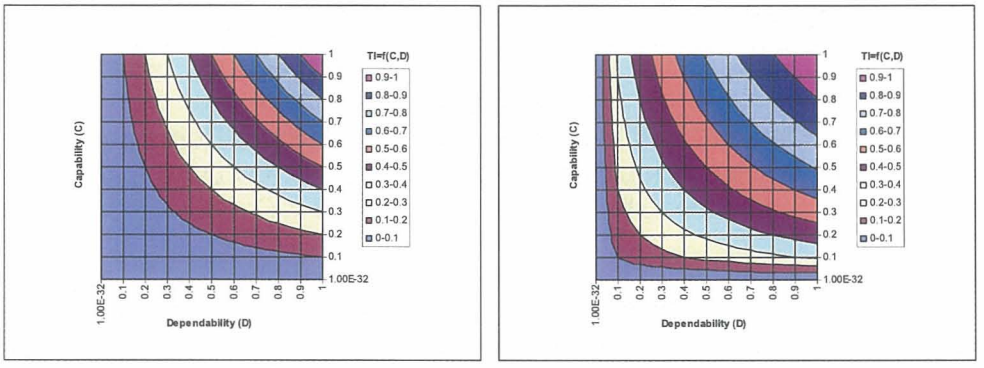


Figure 6. Capability and Dependability with Equal Influence
((i) $\alpha=\beta=1$, $TI = C.D$. (ii) $\alpha=\beta=0.5$, $TI = \sqrt{(C.D)}$)

To confirm the multiplicative conjoint makes sense at the boundaries we consider in Figure 7 the settings for α, β to model extreme conditions where TI is influenced by only one of its attributes.

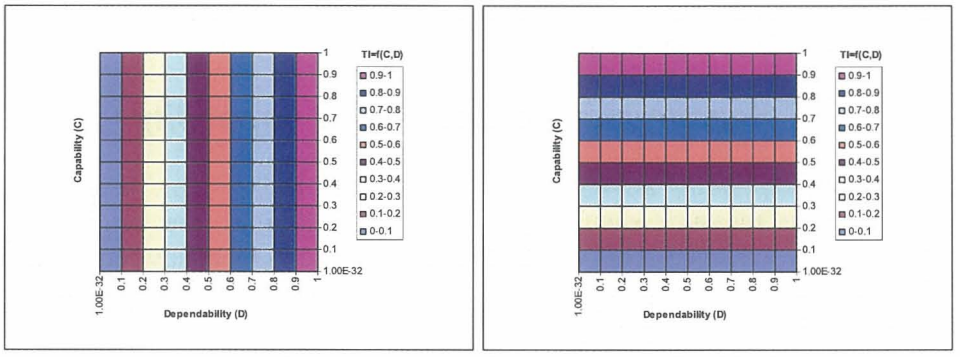


Figure 7. Extreme cases: (i) TI Not influenced by Capability ($\alpha=0$, $TI = D$). (ii) TI Not influenced by Dependability ($\beta=0$, $TI = C$)

The measurement model caters for the more general case where there is a different relative utility of C and D. If this difference in utility is modeled by appropriate selection of values of α, β the shape of the TI function can be skewed for a dependability-biased measure of *technical integrity* or a capability-biased measure of *technical integrity*. This is illustrated for two selections in Figure 8.

[2] A not unreasonable constraint as we shall see when we look at establishing measures for C&D.

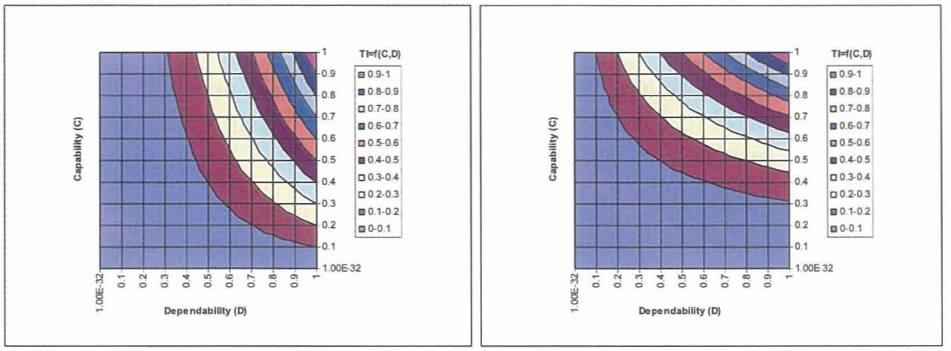


Figure 8. Differing Relative Utilities of Capability and Dependability:
(i) Dependability-biased ($\alpha=1$, $\beta=2$, $TI = C.D^2$).
(ii) Capability-biased ($\alpha=2$, $\beta=1$, $TI = C^2.D$).

Further research is intended to define the approach for establishing α and β empirically with stakeholders. It is noted these utilities may be different for different stakeholders of the system under consideration. An Inter-rater Reliability test will be used to evaluate the chosen values and to establish if consistent values across a range of disparate systems can be arrived at. This would support a finding that the conjoint measurement for *TI* is useful for comparisons across a range of disparate systems.

Notwithstanding the selection of this measurement model, we still have the problem of defining measures for *capability* and *dependability* that may be no more tractable than TI to measure. How are these components to be measured such that the TIs of different systems may be compared? We require either a previously determined measurement scale for each or to determine a simple ordinal scale for each measurand.

A Measurement Approach for Capability

The definition for *capability* arrived at earlier is "that the system meets all specified requirements". The measure should then be the fraction of the originally intended requirements actually achieved. This would equate to failure to meet any requirements being assigned $C = 0$ and all requirements being achieved as $C=1$ and intermediate values measured on the interval (0, 1) as a ratio scale.

While it is common for detailed metrics to be kept about requirements changes and achievements in most system development projects, these may not be readily disclosed to researchers. It may be that for practical purposes, subjective assessments of *capability* may also need to be mapped to values on the scale.

There is precedent for using an ordinal scale relating to requirement achievement. The Study of Systems Engineering Effectiveness (Elm et al. 2007) uses requirements verified as the measure for "Scope", one of the defined elements of "Project Performance" in that study. The study uses a 4 point ordinal scale to represent a subjective assessment of requirement verification. The SE-ROI project (Honour 2010) considers "Technical Quality" as being an ordinal scale where satisfaction of requirements is the "Threshold" level and delivered performance beyond that specified for minimum acceptability, but of value to the stakeholders, is termed the "Objective" level.

A Measurement Approach for Dependability

The definition of *dependability* derived earlier is equated to the absence of risk associated with unintended, undesired emergent behaviour of the designed system is acceptably low and that undesirable limitations resulting from intentional trade-off during design are acceptable. How do we go about measuring the totality of known and unknown risks in a complex system?

Feasibility of Measurement. It seems clear that we should be able to measure *dependability,* at least on an ordinal scale. We should be able to say that System A is more dependable than System B. Measurement on a ratio scale should not be beyond reach. E.g., we can state how much safer (based on historical accident statistics) air travel from City A to City B is compared to a car journey between the same cities.

Required Form of Measurement. Consistent with the structure of the conjoint measurement established earlier, we require a scale $D \in [0..1]$ where 0 indicates the system is not at all dependable and there is certainty in undesired outcomes occurring and 1 indicates the system is 100% dependable and there is no risk of undesired outcomes. Ideally this will be a ratio scale. How to arrive at measurement values for real systems remains a task for future research.

Approaches under Investigation. The literature in the area of measurement of overall risk of unintended consequences of a system largely concentrates on the identification of risks and the qualitative (more often) and quantitative (less often) assessment of each of those risks.

One approach may be to consider dependability as two separately quantified components. One component would be the identified risks and the other the unidentified risks. The identified risks could be combined into a single vale where each consequence identified is associated with a cost (i.e., dollar) value and that value weighted by likelihood of its occurrence. All these weighted values may be summed together to form an overall risk. How to quantify the unidentified risks remains an issue. Some proxy measurement following approaches established for the estimation of faults remaining undetected in complex software intensive systems is possible but arguably not pragmatic to be able to collect across a range of disparate systems. Collection of a range of related measurements and combining them in an assessment tool based on a Bayesian belief network is also possible (Fenton et al. 1998).

Another approach is to look at the number of incidents that breach dependability in the operational life of a system. While this may lead to a more direct form of measurement, there are inherent difficulties that include the infrequent occurrence of high consequence breaches of dependability in critical systems. This makes inference of dependability difficult. A proxy measure where the dependability of the system may be inferred by the more frequent occurrences of low consequence breaches may be a path to further this approach (e.g. (Harris 2004) finds that in aviation the relationship between frequency of occurrence and consequence magnitude for safety incidents follow a power law indicative of self organising critical systems).

A subjective estimating approach for dependability may be to formulate a survey instrument where stakeholders and/or experts in the system may be questioned about dependability characteristics of the system and their responses combined to form a dependability measure. Precedent for this as a pragmatic approach includes the System Usability Scale (Brooke 1996).

Conclusion

A definition of *technical integrity* has been established that combines two attributes of a complex engineered system. The attributes reflect the intended, desired emergent properties of a system and the minimisation of unintended, undesired emergent properties of a system. These attributes have been termed *capability* and *dependability* respectively. It is desirable and feasible to measure *technical integrity* and an approach using a conjoint measurement from measurements of *capability* and *dependability* has been formulated.

Future research is required to further this measurement approach by:

establishing and validating an approach to measure *dependability*. The approach needs to make a pragmatic estimate of the uncertainty of undesired behaviour that is objective, but not too onerous or costly to preclude its use across a wide range of disparate systems.

assessing the relative utilities of *capability* and *dependability* in the multiplicative conjoint measure and seeking to determine if there exists a form of the measurement that may be used across disparate classes of complex engineered systems,

Critical review, ideas and guidance are welcomed by the author to shape this approach into a practical measurement across a broad range of systems to investigate the effectiveness of governance, management and technical processes applied to the development and operation of complex engineered systems.

References

ADF. 2010. Defence Instruction (General) LOG 4-5-012 Regulation of Technical Integrity of Australian Defence Force Materiel. Canberra, ACT, Australia: Department of Defence.

Bale, E.A., and D.W. Edwards. 2000. Technical Integrity - An Engineer's View. *Trans IChemE* Vol 78, Part B (September 2000):355-361.

Barbacci, Mario, Mark H. Klein, Thomas A. Longstaff, and Charles B. Weinstock. 1995. Quality Attributes. Pittsburgh, Pennsylvania: Software Engineering Institute, Carnegie Mellon University.

Boehm, Barry. 1978. Characteristics of Software Quality, TRW Series of Software Technology. New York: American Elsevier.

Brooke, John. 1996. SUS - A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*, edited by P. W. JORDAN, B. THOMAS, B. A. WEERDMEESTER and I. L. McCLELLAND. London.: Taylor & Francis.

Browning, Tyson R., and Eric C. Honour. 2008. Measuring the Life-Cycle Value of Enduring Systems. *Systems Engineering* Vol. 11, No. 3:16.

DoD. 1993. MIL-STD-882C System Safety Program Requirements: Department of Defense, USA.

DoE. 2006. DOE O 413.3A Program and Project Management for the Acquisition of Capital

Assets. Washington, D.C.: US Department of Energy.

———. 2008. DOE G 413.3-1 Managing Design and Construction Using Systems Engineering. Washington, D.C.: US Department of Energy.

Edwards, Michael. 2009. The Impact of Technical Regulation on the Technical Integrity of Complex Engineered Systems. In *INCOSE 2009*. Singapore: International Council On Systems Engineering.

———. 2010. An Ontology of Technical Regulation. In *Systems Engineering Test and Evaluation 2010*. Adelaide, SA, Australia.

Elm, Joseph P., Dennis R. Goldenson, Kaled El Emamm, Nicole Donatelli, Angelica Neisa, and NDIA SE Effectiveness Committee. 2007. A Survey of Systems Engineering Effectiveness - Initial Results: Software Engineering Institute, Carnegie Mellon University and the National Defense Industrial Association.

Fenton, N., B. Littlewood, M. Neil, L. Strigini, A. Sutcliffe, and D. Wright. 1998. Assessing dependability of safety critical systems using diverse evidence. *IEE Proceedinps - Software* Vol. 145, No. I. (February 1998):5.

Gilb, Tom. 2005. Scales of Measure. How to Quantify. In Competitive Engineering. A Handbook for Systems Engineering, Requirements Engineering and Software Engineering using Planguage: Butterworth-Heinemann.

Hackett, Mary. *Technical Integrity and Process Safety*. National Offshore Petroleum Safety Authority 2008 [cited 20 October 2010. Available from http://www.nopsa.gov.au/presentation/.

Harris, David. 2004. Is Aviation Safety a Self-Organising Critical System. Paper read at SETE 2004 Focussing on Project Success, 8-10 November 2004, at Adelaide, SA, Australia.

Hitchins, Derick K. 2003. *Advanced Systems Thinking, Engineering and Management*. Norwood, MA, USA: Artech House.

Honour, Eric C. 2007. Connecting an Architecture with its Emergent Properties. Paper read at Conference on Systems Engineering Research, 2007, March 14-16, at Hoboken, NJ , USA.

Keeney, R.L., and H. Raiffa. 1976. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley & Sons.

Luce, R. Duncan, and John W. Tukey. 1964. Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology* 1 (1):1-27.

Maier, Mark W., and Eberhardt Rechtin. 2009. *The Art of Systems Architecting*. Third Edition ed. Boca Raton, FL: CRC Press, Taylor & Francis Group.

NAVSEA. 2005. NAVSEA Instruction 9410.2 Naval Warfare Systems Certification Policy. Washington Navy Yard, DC: Department of the Navy, Naval Sea Systems Command.

Parnell, Gregory S., Patrick J Driscoll, and Dale L Henderson. 2008. *Decision Making in Systems Engineering and Management*. Edited by A. P. Sage, *Wiley Series in Systems*

Engineering and Management. Hoboken, New Jersey: John Wiley &Sons Inc.

Prasad, Divya Kumari. 1998. Dependable Systems Integration using Measurement Theory and Decision Analysis, Department of Computer Science, University of York.

RAN. 2007. Nias Island Sea King Accident Board of Inquiry Report: Royal Australian Navy.

Roberts, Fred S. 1984. Measurement theory with applications to decision making, utility, and the social sciences. In *Encyclopedia of Mathematics and its Applications*. Cambridge Cambridgeshire and New York, NY, USA: Cambridge University Press.

Sydenham, P.H. 1982. *Handbook of Measurement Science*. Chichester, UK: John Wiley and Sons.

Technical Integrity Audits. PI Asset Integrity Ltd 2007 [cited 20 October 2010. Available from http://www.piai-ltd.com/Technical_Audits.htm.

Tversky, Amos. 1967. A general theory of polynomial conjoint measurement. *Journal of Mathematical Psychology* 4 (1):1-20.

Biography

Michael Edwards holds a Bachelor of Engineering (Honours) and a Master of Engineering Science (Communications) from the University of NSW and a Master of Business Administration (Technology Management) from Deakin University. He is an Engineering Fellow with Raytheon Australia and is a part-time research student at the Defence and Systems Institute of the University of South Australia. He is a Fellow of the Institution of Engineers, Australia, a Chartered Professional Engineer, a member of the Systems Engineering Society of Australia and a Member of INCOSE. His professional and academic interests build on an extensive experience base in the practice of engineering management and systems engineering largely in the domain of complex naval systems including applications in surface combatant and submarine combat systems, mine warfare systems, hydrographic systems and naval aviation systems.