**Revie, M. and Bedford, T. and Walls, L. and Shimell, J. and Baldwin, T. (2016) What is the value of a standard? In: Risk, Reliability and Safety. CRC/Taylor & Francis Group. ISBN 9781315374987 ,**

# What is the Value of a Standard?

M. Revie, T. Bedford & L. Walls
Department of Management Science, University of Strathclyde, Glasgow, Scotland

J. Shimell & T. Baldwin
Polaris Consulting, UK

ABSTRACT: Standards play a critical role in the procurement of defence, and other, systems. Choosing the most appropriate standard is important but has become more topical given the UK Ministry of Defence policy of "as civilian as possible, as military as necessary". Whereas historically managers might have selected from classes of defence standards, this choice set is now increased to include civil standards. We develop a model that has been commissioned by the UK Defence Standardisation whose responsibilities include supporting project teams on the selection of standards. Our model is based on an extension of Bayesian Belief Networks, called an Influence Diagram, which allows decisions and consequences to be represented as well as uncertainties. We have developed an initial model for a real case to assess the feasibility and use. We outline the context of the defence procurement project in our case study and describe the reasoning underpinning the model structure. We have found that it is possible to develop a simple model that captures the views of multiple stakeholders and informs a reasoned choice about the value of alternative standards.

## 1 INTRODUCTION

The UK Defence Standardization (DStan) is the Ministry of Defence's (MOD) centre for "throughlife standardization and its management across defence" [https://www.contracts.mod.uk/defence-standards/]. Currently DStan manages a portfolio of circa 900 standards. As well as maintaining an up-to-date relevant portfolio through the addition, revision and removal of standards, DStan also offers other services, such as advice and support to project teams on the selection of standards within activities throughout the Concept, Assessment, Demonstration, Manufacture, In-service and Disposal (CADMID) cycle. For more details about the responsibilities of DStan please see https://www.gov.uk/guidance/uk-defence-standardization.

DStan contributes to standardization policy-setting for international defence organisations, such as NATO and the European Defence Association (EDA). More generally, MOD staff participate in committees responsible for the development and revisions of civil standards, such as those maintained in the UK by the British Standards Institute (BSI). These will include dependability (i.e. reliability, availability, maintenance and maintainability) but the scope of standards is broad and includes management processes (e.g. for safety) and technology standards (e.g. materials).

Recently DStan has adopted a policy towards standardization which it labels "as civil as possible, as military as necessary". See, for example, [https://www.contracts.mod.uk/announcements/uk-defence-standardization-dstan/] for details of a call to civil standards organisations to apply for a government license in this respect. This policy has emerged in response to the changes in structure and composition of the UK MOD and the Defence Industry, as well as in light of evolving technological challenges and opportunities. Thus, the MOD are increasingly considering the use of civilian standards where possible and only use Defence Standards when absolutely necessary. This raises challenges in establishing the suitability of civil standards in a military environment. There are examples such as management standards which might be transferable, but there are other examples such as for some technologies where the usage stress levels and profiles might differ substantially between civil and military use.

This policy also initiates change within the DStan organisation itself since its historical role needs to adapt to managing this larger and more complex standards portfolio. In particular, the nature of the advice provided to project teams is changing since not only is it important to select appropriate standards for contracts, tests and so on, but the consequences of using an inappropriate standard can be costly. Thus, DStan is in transition from an organization that viewed itself to a large degree as a repository of knowledge and information specific to the military context, to positioning itself as a service provider. However, understanding the value of the service that DStan may provide, is inherently linked to the added value of the standards they offer. Demonstrating this added value is very challenging.

To support a standards organization like DStan, a rigorous and defendable method is required to help users select appropriate standards. The purpose of such a method would be to enable project managers to become intelligent users of standards and identify the most appropriate standard (if any) to adopt on a given project.

The aim of this paper is to develop a method through which those users selecting standards for a particular project purpose can assess the trade-offs they are making when selecting one candidate standard over another. For example, candidate standards might be a civil standard, a modified civil standard (e.g. with additional annex for military environment), or a defence standard). Different standards will typically provide different levels of confidence about the performance of the system under operational conditions, and hence the decision for one standard has similarities to a "value of information" decision problem.

To achieve our aim, we have identified and developed a method based on a graphical decision model to help users understand the value of different standards on a given project to procure a new defence product.

In the remainder of this paper, we position our work relative to the wider literature on standardization in Section 2. In Section 3 we describe the principles of our method, while in Section 4 we describe a case study to evaluate and develop the proposed method and discuss what we have learnt through this implementation. The case study is real although details have been de-sensitised in our description of the model developed. Future developments of the modelling tool and additional research required are discussed in Section 5.

## 2 LITERATURE REVIEW

There appears to be no literature that addresses the problem posed in this paper; that is modelling to provide guidance on the adoption and selection of standards in general to meet the purpose of the specific use application. There are journal articles that refer to the relevance and use of particular standards for a specific purpose. For example, Leitch (2009), Purdy (2010), Aven (2011) in relation to ISO 31000 for risk management and Braband and Griebel (2004) in relation to IEC 61508 for functional safety. Such papers tend to examine a particular standard from through either the lens of its core principles or by reflecting upon experience of implementation within a particular organization setting. The challenge addressed in this papers includes standards for risk, safety, reliability and maintenance but is not restricted to these domains.

More generally, Timmermann and Epstein, (2010) have observed that that since standards tend to be defined for a wide community and are hence generic, then challenges emerge for organizations as they manage adopt within their particular context. This has implications for our context since defence standards are developed to reflect UK military use and are widely used by industry when developing products or services for the MOD. Therefore if the MOD make greater use of civil standards, which are not designed specifically for military use then they may not meet MOD requirements. This implies that either the MOD accepts this risk or places additional effort tailoring civil standards to bring them to a level that is suitable for the defence context

Historically, standards have been viewed as limiting innovation in products and processes. For example, Arthur (1989) and Katz and Shapiro (1992) all discuss the role of standards in restricting innovation. However, more recently greater consideration has been given to understanding how standards can be and have been used to promote innovation. For example, Blind (2013), Swann (2000), Swann (2010), and Narayann and Chen (2012) discuss how innovation has been promoted and facilitated through the use of standards, although they also recognize the dis-benefits of standardization in their attempts to provide a balanced, critical analysis. These papers, and related work of the authors, have been based data collected from panel surveys, company cases, literature reviews, economic modelling and sociological analysis. Amongst the reported positive influences of standardization on innovation are the following; adverse selection of options; economies of scale; creating environment of trust; efficiency in supply chains; and reducing transactional costs. While negative effect on innovation from the use of standards include: raising rival's costs; reducing choice; and prematurely selecting technology solutions before markets mature. Recent conference papers by Tasker et al., (2014a,b) specifically discuss the need to encourage innovation in within the framework of through-life engineering services, a context relevant to the MOD and the defence industry, through the development and use of service standards. That is, Tasker et al (2014a,b) propose moving away from technical standards on materials, to develop framework standards on how to use equipment or create value from equipment. While we agree that process based standards, including those focusing on the integration of the engineering products and services, are important, we do not agree that there should be a move away from technical standards. To the contrary, many technical standards are vital to support the development of safe and reliable products that can be maintained through life. Indeed it is particularly the technical standards that are of most interest from our modelling perspective since these are the ones in which the differences between the military and civil contexts are important risk sources and hence where choices need to be made in their selection and implementation.

## 3 MODELLING METHODOLOGY

This section describes the methodology adopted for our decision model to evaluate the impact of selecting candidate standards for a defence procurement project. Throughout we refer to the procurement of systems and we use this term to encompass everything from small pieces of equipment to subassemblies and complex systems.

### 3.1 Model requirements

The model selection criteria have been developed in collaboration with MOD Requirements Managers. These are the managers who are responsible for selecting the standards to be used during the system procurement process. The Requirements Managers have relevant experience applicable to a range of typical defence procurement projects and so we believe that the criteria defined are generalizable within our context.

Through discussions with the Requirements Managers we have been able to articulate the challenges affecting the selection of a suitable standard during defence system procurement. These challenges are as follows. First, understanding the dependency between the overall system requirement, the standard selection and the successful outcome of the procurement project, typically defined as timely, affordable and to specification. Second, articulating the different consequences of selecting a standard in relation to the desired outcome. Third, understanding the impact of observing sources of data on the likelihood of being able to procure an effective system. Finally, capturing the uncertainty inherent in all aspects of selecting a standard and so procuring the system.

Our modelling methodology must be capable of addressing these challenges. In particular, our chosen model type should be capable of the following in relation to the third and fourth challenges. First, capturing uncertainty and dependencies between random variables. Second, estimating the likelihood of a parameter of interest being in a given state based on all available data and knowledge. In addition, through discussion with the Requirements Managers it was surfaced that it would be beneficial if the modelling methodology support engagement with managers and other stakeholders across a variety of project contexts and types of standards. This suggests that being able to create a graphical representation of the decision problem to provide an intuitive way to visualise dependencies and to structure a problem would be desirable. It is also important that our modelling method be comprehendible and usable by the practitioners who regularly use and apply standards.

### 3.2 Bayesian belief network model class

Following the above considerations, we decided that the Bayesian Belief Network (BBN) modelling approach should be evaluated as the basis of our methodology. This is because BBNs can be highly effective means of modelling to reason through relationships which are highly uncertain and this model class has all the characteristics described above to fulfil the Requirements Managers criteria.

Determining the value of a standard within a procurement project is complex and likely to require multiple dependent factors to be modelled in a structured and transparent manner. Further, BBN and their natural extensions called Influence Diagrams (ID) allow us to formally represent not only the uncertainties associated with random variables, but also to explicate the decision variables (e.g. which standard?) and the valuations (e.g. consequences of selecting a given standard to achieve purpose).

For brevity, the theoretical foundations of BBNs and IDs are not discussed here. Pearl (1998) provides a foundational text, while Fenton and Neil (2012) provide descriptions of the principles and the applications of BBN and ID models.

In summary, the structure of a BBN comprises of two key elements; nodes and arrows. Nodes represent uncertain variables while the arrows represent causal or influential links between nodes. The direction of the arrow represents the direction of influence. Each node and arrow contains probabilistic information, either as marginal probability distributions or conditional probability distributions. Typically the term BBN is used for models that only contains random nodes. A BBN can be generalized to be an ID model by including decision nodes and consequence nodes. Decision nodes are represented by a square and where included in the model, they can be used to identify the optimal decision. Consequence (or value or utility) is represented by a hexagon. These nodes capture the value of different outcomes and are used to model aspects such as cost of testing, whole life cost, etc.

BBN/IDs have features that make them appealing for assessing the value of different standards. First, the BBN/ID uses graphical tools to qualitatively structure the problem. Those who regularly use BBNs and IDs, or other similar visualizations, believe that it promotes effective communication between decision makers and experts, and that it improves the credibility of the model. Second, BBNs/IDs are based on strong mathematical foundations providing a theoretically robust and consistent method of combining multiple sources of information. Third, future decision makers will be able to understand the justification for previous decisions if the qualitative and quantitative information has been elicited in a transparent manner. This was raised as an important feature during our initial discussions with practitioners.

### 3.3 Model building process

One potential limitation with adopting a BBN/ID approach is the resource required to develop the model. Stakeholder involvement and support is required to develop a meaningful model and to populate it with, for example, probabilistic data, valuations as well as defining the set of decisions. However the level of involvement required of stakeholders will depend upon the complexity of the problem being explored and the availability of data.

To address this potential limitation and ensure resource demands on project teams involved in model development are reasonable, we have designed a one day workshop with relevant stakeholders to structure the ID model, This workshop has been implemented as part of a case study to build an initial qualitative model to evaluate the value of different standards for a particular piece of equipment for which this was a topical issue.

In terms of system scale, the equipment of interest for the case study is at the less complex end although it is a vital item for which extensive testing was required to military standards. The choice of a simpler system for the initial case has been deliberate to develop a prototype model that can be challenged and adapted for a wider class of system in due course.

## 4 CASE STUDY

The aim of the case study is twofold. Firstly, to evaluate the modelling methodology in terms of its potential as a tool to support MOD Requirements Managers understand the impact of selecting different standards. Secondly, to help structure the knowledge and thinking of the Requirements Manager through the process of developing the ID model and so qualitatively support his/her decision making process.

### 4.1 System characteristics and context

The case study focused on the procurement of a simple system – analogous to personal protective equipment. In procuring the new generation of this type of system, the MOD have the opportunity to buy and use an alternative system from the system family currently in use. However, as the alternative system was tested to different standards than the UK MOD norm, there was uncertainty regarding the consequences of adopting the alternative equipment.

For this particular case study, the project team sought to procure new equipment and define what standards to use as part of their requirements definition process. The project team were reluctant to reuse previous standards without fully understanding the consequences, especially when alternative standards or options were available to them.

The focus of the case study has been to develop a modelling structure to allow the Requirements Manager to understand the impact of selecting different standards. The structuring was facilitated during the one day workshop with members of the project team. Namely the Requirements Manager and two Subject Matter Experts (SME).

### 4.2 Workshop design

The workshop was facilitated by two of the authors and was split into two parts. During the morning session, which lasted approximately three hours, the two SME and the Requirements Manager discussed the differences between each of the standards options, including what they impacted upon. The facilitators began this process by understanding the different decisions available to the Requirements Manager. This was followed by understanding the consequence of these decisions. Through this, the different valuations, and hence expression of utilities, and the uncertainties that impact these utilities were identified.

Following the morning session, the facilitators structured the information gathered into an ID model. During the afternoon session the modelling structure was explained and the SMEs validated the structure. The different variables were explained together with the structure of the network. Based on feedback, modifications were made to the structure where appropriate.

Here we discuss the three different types of variables in the model and what they represent and influence.

### 4.3 Decisions

During the workshop discussions, two different decisions available to the Requirements Manager were identified. The Requirements Manager must choose whether or not to use a standard, and if so, which standard to adopt. In addition, the Requirements Manager is responsible for determining the system requirements, and can influence how strict these are. These decisions must be made in parallel as the Requirements Manager does not want to select a system requirement that is incompatible or conflicts with a chosen standard. Thus the model includes the two decision nodes; Choice of Standard, and System Requirement. For this case, this included a Defence Standard, two UK Civil standards and no standard.

### 4.4 Uncertainties

The project team believed that there were a number of factors influenced by the choice of standards and the system requirement. Firstly, they believed that the different choices of standard available to them influenced factors related to the test. This is because alternative standards proposed more (or less) tests to be carried out and passed to demonstrate conformity. Secondly, the tests had different characteristics, such as stress levels and profiles, them that gave the team varying degrees of confidence that the system would be tested under conditions that allowed them to make reasonable inferences about performance under representative operational use conditions.

Differences between the standards included type of pre-conditioning (e.g. none, some), degree of damage (e.g. functioning, failure) and the allowable inactive period between consecutive tests (e.g. the duration between shocks imposed to stress the system). Depending upon the states of these three variables, the team believed it was more or less likely that the test was representative of operational conditions.

From this discussion five different uncertain variables were created. Namely: Number of Tests; Pre-Conditioning; Degree of Damage; Period between Tests; and Operational representativeness.

The team also believed that the choice of standards would impact on the probability of the system being interoperable, Interoperability has been interpreted very broadly as the ability of the equipment to be used in conjunction n within both UK forces and international coalition partners. Thus interoperability is important for military systems given the nature of their operational use. Military standards were considered more likely to produce interoperable equipment when compared with the options of using civil standards or no standard.

The team also stated that they believed the likelihood of the system being exported to other countries was impacted by its interoperability. For example, a system that has been designed to a NATO standard was believed to offer greater potential to a company looking to export to other NATO nations than a system designed to a UK Defence Standard.

From the discussions around these two aspects of interoperability, two new random variables were created. Namely: Interoperability; and Exportability.

The team further believed that the strictness of the system requirement would impact on the operational advantage gained by the MOD on using the system. The strictness of the system requirements would also impact on the probability that the system meets the requirements as documented in the so-called System Requirements Document (SRD); with more challenging requirements being less likely to be met. Whether or not the equipment meets the SRD clearly influences the military advantage the MOD gains from use of the system. From this discussion, two new variables were created. Namely: Operational Advantage; and Equipment meets SRD. Note that Equipment meets SRD can be interpreted broadly. The SRD documents all the requirements of the system, including the system being able to meet performance and safety requirements. All these features may be included within the variable Equipment meets SRD. The team believed requirements criterion with priority in this case was safety. However, in other cases, this variable may require splitting into multiple variables to capture the multiple criteria in relation to performance stated in the SRD. Further, Operational Advantage was viewed as the system being capable of meeting the current military capabilities.

Finally, the team were interested in the output of the tests as this is the only variable which can be truly observed. A new variable was created; Test Results, which was influenced by the following three variables: Number of Tests; Representative with respect to in-service use; and Equipment meets SRD. The logic behind this structure was as follows. If the equipment meets the requirement, it is more likely to pass the test. If the test is unrepresentative of operational conditions, then the test is unlikely to identify systems that do not meet the requirement. Finally, having a small number of tests is also less likely to fail poor systems.

### 4.5 Consequences

The team believed that ultimately the decisions and uncertainties previously described would impact four different financial metrics. These metric represent the consequences that are to be valued and hence represented as the utilities within the model.

Firstly, the choice of standard would influence the upfront or initial procurement costs. Different aspects of the test would cost more; e.g. smaller tolerable levels require more expensive raw material and manufacturing methods. Having a larger number of tests to carry out would also increase the up-front costs.

Secondly, the team believed that if the system was placed in-service (since it passed the test), but it was subsequently found to not meet the original requirements or be sufficiently interoperable, then rework would have to be carried out. As such, rework costs would be incurred in either of these scenarios.

Thirdly, the team believed that exportability would potentially offer the MOD the opportunity to recoup costs via export levies imposed on any industry systems that were sold to non-UK parties. If the system was exportable, then costs could be recovered in the future.

Fourthly, the summation of the up-front costs, the reworks costs and the export levy would give a whole life cost. However, this is not the whole life cost of the system, but rather the difference in the whole life cost from adopting the alternative standards.

Therefore, four different consequence nodes were created. Namely: Testing or Procurement Costs; Rework Costs; Export Levy; and Whole Life Cost.

### 4.6 Decision model for selecting between standards

Figure 1, on the following page, shows the ID model developed to capture the relationships between the uncertainties, the decision nodes and the consequences discussed above.

### 4.7 Implications of modelling process

The process of structuring the model encouraged a number of issues to be discussed by the team during the workshop. In particular, the following issues were revealed through the modelling process.

Firstly, while the alternative standards contain a considerable amounts of information, we have found that the core aspects that differentiate between different types of candidate standards in terms of content and their subsequent use within a decision making process can be distilled down to a relatively small number of key variables.

Secondly, the representativeness of the test was a key variable to the SMEs and was important in determining how much value they attributed to a given type of standard. In particular, there were differences of opinion within the team about the importance of certain aspects. For example, the impact of preconditioning on the equipment functionality.

Thirdly, the team agreed that the standard could be viewed as an insurance policy. That is, the defence standard may lead to greater costs that a civil standard. But a defence standard also provides a greater level of reassurance regarding the functionality of the system.

Finally, the process also raised the question – "Do we need to quantify the model?" For example, is the learning gathered through the process about the key issues sufficient for the Requirements Manager to make a decision without quantifying the model? Could the model be used in a qualitative manner in some cases? These questions have not been explored within the case study to date. But we suggest that this should be considered prior to any model quantification process.

## 5 CONCLUDING DISCUSSION

### 5.1 Future use of the model

The model has been developed in order to support the Requirements Manager in his/her decision making. As argued above, we believe that the qualitative model building process in itself helps the Requirements Manager think through the reasoning for different choices of standards. If quantified, the model could also provide numerical output that might further allow comparison of options through answering questions such as the following.

- Which standard should be used and why?
- Which standard is the 'best' option for meeting the system requirements?
- What defines 'best' – is it value-for-money or maximise capability or minimise risk or minimise cost or maximise commercial opportunities?
- Assuming that a given standard and a system passes the standards test, how likely is it that the system does not meet MOD requirements?
- What are the implications with respect to standards of choosing strict system requirements?
- Under what circumstances would Standard X be a better choice than Standard Y?
- How robust is Standard X to changes in future operating scenario?
- What would be the expected cost of adopting certain standards if some of uncertainties become fixed? For example, the system must be interoperable.

### 5.2 Main conclusions

Our main findings are as follows.

We have developed a novel approach to assessing the value of a standard in a practical context. Our approach is distinctive not only because of our context but also because it differs from the existing literature on standardization which tends to be dominated by principled assessments and practical reflections about specific standards, socio-economic implications of standardization more generally especially in relation to product and process innovation.

We have found that BBN/ID models have features that make them appealing for assessing the value of alternative standards. The use of graphical methods to support the qualitative structuring of the problem, enabling project teams to understand the impact of standard selection and for this problem structure to be shared quickly. This class of models is based upon strong mathematical foundations providing a theoretically robust and consistent method of combining multiple sources of information from different project team members. Such decision models encourage the transparent elicitation of.

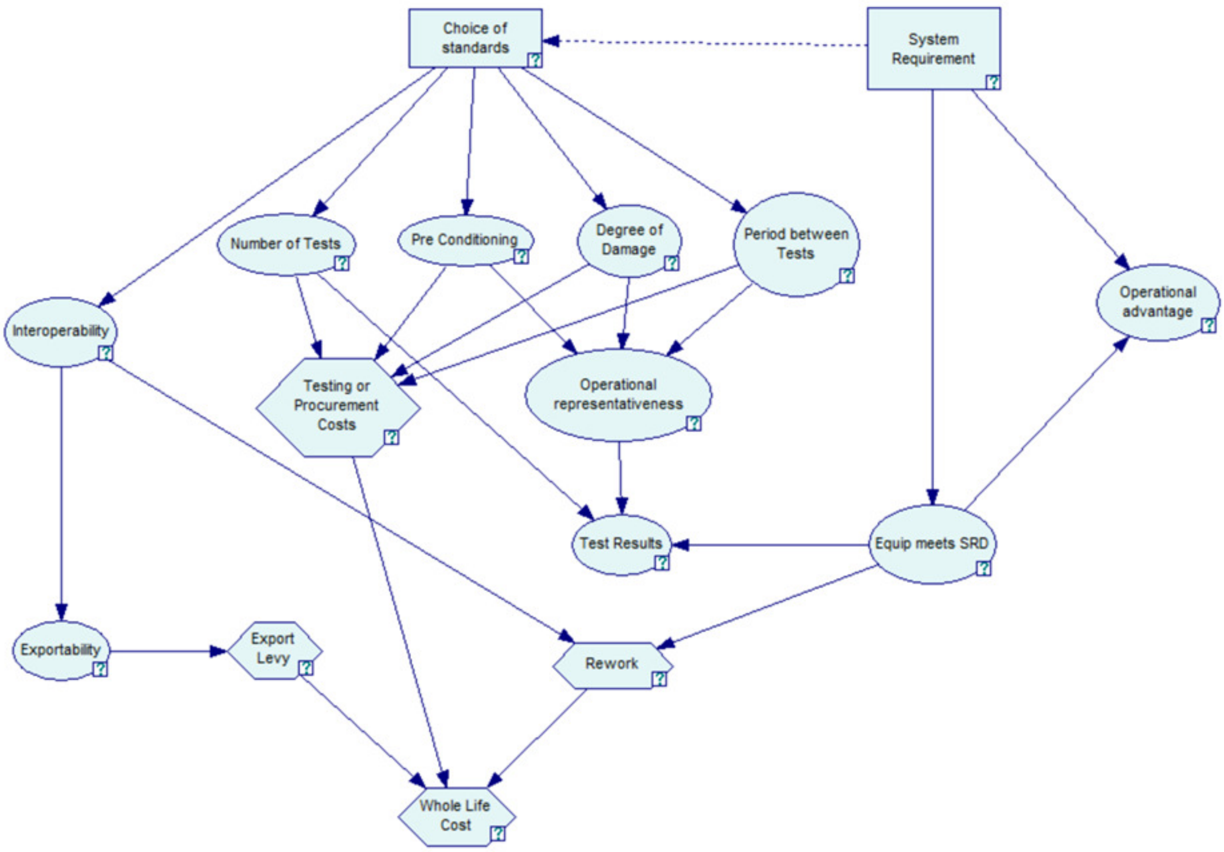


Figure 1. Decision model to assess value of a standard developed for defence case study

qualitative and quantitative data and information. Importantly, future decision-makers should be able to understand the justification for previous decisions, i.e. why a particular standard was selected.

Our case study has demonstrated that it is feasible to gather the necessary qualitative information to structure a model for assessing the value of standards. Indeed the decision facing the project team involved in the case was real and the model did provide useful feedback to the Requirements Manager.

The process of structuring the model in our case study has identified key features of the selection process. Experience through the case suggests that often a decision-maker can gain sufficient insight to make a choice based on a qualitative reasoning and hence implying that a full quantitative model is not necessarily required. In situations where the difference between alternative standards is considered to be negligible, then the added resource required to quantify the model may mean that in modelling terms, the options are equally good.

### 5.3 Further Work

There is a need to explore the potential for using this type of model more generally. Our case study has proven in principle the feasibility of the method.

But as we acknowledged, our system under consideration is comparatively small and self-contained. More generally there is a need to evaluate the use of the model for a multi-dimensional case study (i.e. multiple interacting systems, multiple types of standard). This would allow the relationships between the different dimensions to be explored and so determine if our proposed approach remains suitable.

There is also a need for a proper evaluation of the added benefits of developing a fully quantified model in relation to a qualitative model that captures the reasoning for less resource to build.

## ACKNOWLEDGEMENTS

We would like to acknowledge the contribution of all the MOD staff that took part in conversations, workshops and provided feedback on the model development.

## REFERENCES

Aven T. On the new ISO guide on risk management terminology. Reliability engineering & System safety. 2011;96(7):719-26.

Blind, K., The impact of standardization and standards on innovation. Report within the MIoIR-

NESTA Compendium of Evidence on Innovation Policy. London/Manchester, 2013.

Braband J, Griebel S, editors. Safety Analysis according to IEC 61508—Putting it into Practice. Probabilistic Safety Assessment and Management; 2004: Springer.

Fenton N, Neil M. Risk assessment and decision analysis with Bayesian networks: CRC Press; 2012.

Leitch M. ISO 31000: 2009—The new international standard on risk management. Risk Analysis. 2010;30(6):887-92.

Purdy G. ISO 31000: 2009—setting a new standard for risk management. Risk analysis. 2010;30(6):881-6.

Tasker, P., A. Shaw, and S. Kelly, Through-life engineering services–innovation and the role of Standards. Procedia CIRP, 2014. 22: p. 197-203.

Tasker, P., A. Shaw, and S. Kelly, Standards for Engineering Services. Procedia CIRP, 2014. 22: p. 186-190.

Timmermans, S. and S. Epstein, A world of standards but not a standard world: toward a sociology of standards and standardization*. Annual review of Sociology, 2010. 36: p. 69-89.

Swann, P., The economics of standardization: an update. Report for the UK Department of Business, Innovation and Skills (BIS), 2010.

Swann, G.P., The economics of standardization. University of Manchester, Manchester, UK, 2000.

Narayanan, V. and T. Chen, Research on technology standards: Accomplishment and challenges. Research Policy, 2012. 41(8): p. 1375-1406.

Arthur, W.B., Competing technologies, increasing returns, and lock-in by historical events. The economic journal, 1989: p. 116-131.

Katz, M.L. and C. Shapiro, Product introduction with network externalities. The Journal of Industrial Economics, 1992: p. 55-83.

Pearl, J., Probabilistic reasoning in intelligent systems: networks of plausible inference. Series in representation and reasoning. 1988, San Fransico, CA: Morgan Kaufmann.