28th annual INCOSE international symposium
Washington, DC, USA
July 7 - 12, 2018

Use of TRL in the systems engineering toolbox

Kjersti Bakke
University College of SE Norway
Department of Science and Industry Systems
kjersti.bakke@gmail.com

Cecilia Haskins
NTNU
Mechanical & Industrial Engineering
cecilia.haskins@ntnu.no

Abstract. Technology Readiness Level (TRL) is used to evaluate the maturity of the Critical Technology Elements (CTE) in a project or program and indicates an assumed level of the remaining technical risks for each independent technology. This paper presents results of a literature review exploring the diversity of definitions applied to TRL. Survey results explore whether and how TRLs are used. Based on these results the authors assess to what degree TRL evaluations are suitable for organizations without separately trained and skilled readiness assessment personnel.

Introduction

The technology readiness scale is to some degree mystical from an engineering viewpoint, the scale is ordinal in nature, and a level is based upon reaching a technical milestone, such as e.g. a simulated test of a prototype. Meters and other measurements normally applied in engineering are strictly logical, and have a set internal relations or external to other units. This thesis will examine why and how the TRL scale works.

Background

History: TRL is a system engineering tool that was created to perform maturity evaluation of technology. The metric was developed by Stan Sadin, NASA, and was formalized after years of testing (Sadin, et al., 1989).By defining delivery quality as a product of cost and schedule; the TRL scale soon showed to be very useful in assessing the expected delivery quality. TRL was originally created to achieve a "mutual agreement between research personnel, research management, and mission flight program managers" (Sadin, et al., 1989). The intention was to differentiate technology maturity in a disciplined independent way. The original definition of TRL was a scale from 1; that could be a napkin-model defining basic ideas or concepts – to 7; a system prototype demonstration in space environment (Sadin, et al., 1989). The method was later expanded to include up to 9 levels (Mankins, 1995) which is the currently prevalent model. An additional 10th and 11th level have been recommended in order to make the technology readiness assessment (TRA) commercially available (Straub, 2015; Hicks, et al., 2009). Most recently, Austin et al. (2017) described application of Bayesian networks to the TRA.

Suitability: The TRL scale is used as an indicator of the embedded technical risk of a project. Due to the limitations of the measurement scale, several factors that also influence risk are not a part of the TRA. Nonetheless, TRLs remain a part of basis for decision making in projects and programs.

It is evident from the persistent use of TRLs that the scale has some elements of usefulness, which begs the question why and how it works. How do people understand this maturity measurement? Under what conditions is it used or not used? From a systems engineering viewpoint, how does it affect the way people work within projects?

Research questions: The unanswered questions resulted in these research questions:

- TRL can be stated in requirements, does this drive the need for systems engineering?
- TRL can be used to determine risk, does this drive the need for systems engineering?
- TRL often contributes to determination of budget and schedule parameter, does this drive the need for systems engineering?

Research Methods

The overarching framework for this inquiry is the technology readiness assessment (TRA) and resulting TRLs. A methodological fit may be considered as an "internal consistency among elements of a research project" (Edmondson et al., 2007). The research questions were addressed by considering the topic from the following angles.

A literature review explored previous publications with regard to the use of TRLs. Both positive and negative experiences have been reported, and were included. For possible replication of this work; the sources used for the research include online sources such as: Science Direct, Elsevier, ResearchGate, Google scholar, and government/organization webpages. Search phrases used included Technology readiness levels, TRL, Technology readiness assessment, TRA, new product development, and specific reference authors' names.

A content analysis was performed to compare the information gathered about the different levels and how they are understood both in research literature and in user organizations. For this reason a large number of the references point to user organizations. The 9-level TRL scale is used as a basis for the content analysis. All 7-level scales have been excluded from the content analysis to avoid noise.

A Survey was conducted in order to see how persons working in new product development or other activities related to systems engineering think about the TRLs. Survey response were collected trying to pinpoint:

- The intuitive understanding of the different TRL levels
- The knowledge about the use of TRLs
- The extent to which the respondents thought TRL or another form of maturity metric may be suited as a basis for evaluations that influence the progress of a project or program.

The survey was distributed among personnel with experience working with new product development, and a link was posted on the INCOSE Official site on LinkedIn. The respondents were asked to avoid looking up the TRL scale before they completed the survey. The survey was oriented toward technical personnel and technical management; at all levels from assembly to portfolio managers. The results presented have been filtered so that answers that both confirmed and repudiated the same statement or were hastily filled with the same answer throughout the survey were removed. The questions were prepared in Norwegian and English. After the survey was collected two questions were removed as the translation was deemed too weak to yield reliable results.

Limitations in the literature reviews relate to the source of information. Many of the existing papers on TRLs are based on a limited number of programs run by US Government Agencies. Several of the authors have used numbers directly from NASA's "Resource Data Storage and Retrieval data base" (REDSTAR) for their studies or used the Lee & Thomas (2001) study as a basis (Gatian, 2015; Dubos, et al., 2008; Dubos et al., 2011; Engel, et al., 2012; Azizian, et al., 2009; Katz, et al., 2015; Meier, 2008) . The knowledge captured in the above-mentioned studies are likely to be colored by the environment in which the information was gathered, but may still give an indication of the value of TRL that could be applicable also in other domains.

Theoretical background

The TRL level of the Critical Technology Elements (CTE) in a project or program indicates an assumed level of the remaining technical risks for each independent technology. In guides, this is referred to indirectly as a “development hurdle” (ESA, 2008; Mankins, 2009) or “maturity gap for successful inclusion” (US GAO, 1999). The TRA is a tool created to overcome the difficulty of comparing discipline-specific metrics when measuring the feasibility of products across disciplines (Sadin, et al., 1989). The technical maturity on a discipline independent level can be vital to evaluating whether or not a technology is mature enough to proceed to integration. The US Government Accountability Office (1999) has stated that: *«Our best practices work has shown that a technology readiness level of 7— demonstration of a technology in a realistic environment—is the level of technology maturity that constitutes a low risk for starting a product development program.”* In the quote GAO connects TRL and residual risk that will be carried into the program or project. US Department of Defense (DoD) have considered a TRL equal to or greater than level 6 to be a mature technology; ref (Conrow, 2009), the Science technology objective (STO), among other matures technology up to TRL 6 for further implementation in US. DoD programs (Graettinger, et al., 2002). Meanwhile DoD have agreed with GAO that a TRL 7 is preferable upon integration in programs already in 1999.

Organizing delivery around TRL. Dacus (2012) concluded that by using squared TRL shortfall (eq. 2) to communicate the relation between the ‘TRL shortfall at integration’ and the ‘severity of the additional risks by integrating immature technology’; risks may be defined before integration. Dacus further presents the following equations where TRL_i is the TRL level of the technologies to be integrated;

Weighted severity of issues $SS = \sum_{i=1}^{N}(8 - TRL_i)^2$ (eq.1)

Maximum technical shortfall $Max = Max(8 - TRL_i)$ (eq.2)

The Dacus study further applies these equations to past projects, and shows that with less than the recommended *3 maximum technology shortfalls* have a mean of 7,7 months’ delay, with additional costs of 3,2 %, while the projects that violate this rule had a mean 31,2 months’ delay and a 35,5 % cost overrun.

TRL Heterogeneity. (Dubos, et al., 2011) present result of integrating two sets of technologies with average TRL at μ_{TRL}=6. One set shows heterogeneity of TRLs, the other has a variable set of TRLs for the technology. The study concludes that a homogeneous TRL pool with all TRLs included equal to TRL 6, is clearly beneficial for the overall schedule risk, as compared to a variation of TRL levels with similar average

To increase understanding of the TRL scales inherent nature, a weighted TRL scale was presented by Conrow (2009). The TRL scale is originally an ordinal scale where each step has a different inherent value. Conrow has used Analytic Hierarchy Process (AHP) to transition from an ordinal to a cardinal scale, distributing the steps according to resource use; the results from (Dubos, et al. 2011) have been reassessed below to see to what degree the average TRL is in fact still an average. The TRL values of the arrays (pf_{1D} and pf_{2D}) were exchanged with the equivalents of Conrow (pf_{1C} and pf_{2C}, to see if the scales ordinal character is the reason for the effect found in the study (Dubos, et al., 2011);

$pf_{1D} = [6\ 6\ 6\ 6\ 6\ 6\ ]$ and $pf_{2D} = [4\ 5\ 5\ 7\ 7\ 8\ ]$ where $\overline{pf_{1D}} = \overline{pf_{2D}} = 6$

The resulting arrays when converted to Conrow’s weighted TRL scale are

$pf_{1C} = [2{,}74\ 2{,}74\ 2{,}74\ 2{,}74\ 2{,}74\ 2{,}74\ ]$ and $pf_{2C} = [1{,}14\ 1{,}97\ 1{,}97\ 4{,}26\ 4{,}26\ 6{,}81\ ]$

Averaged variance change from

$$\delta_{1D} = 0 \quad and \quad \delta_{2D} = 1{,}42 \ \text{ changed to } \ \delta_{1C} = 0 \quad and \quad \delta_{2C} = \ 0{,}56$$

Corrected for the weighted scale, the average of the distributions change as shown below

$$\overline{pf_{1C}} \ = 2{,}7 \quad \text{and} \quad \overline{pf_{2c}} \ = 3{,}4 \quad \Rightarrow \quad \overline{pf_{1C}} \neq \overline{pf_{2C}}$$

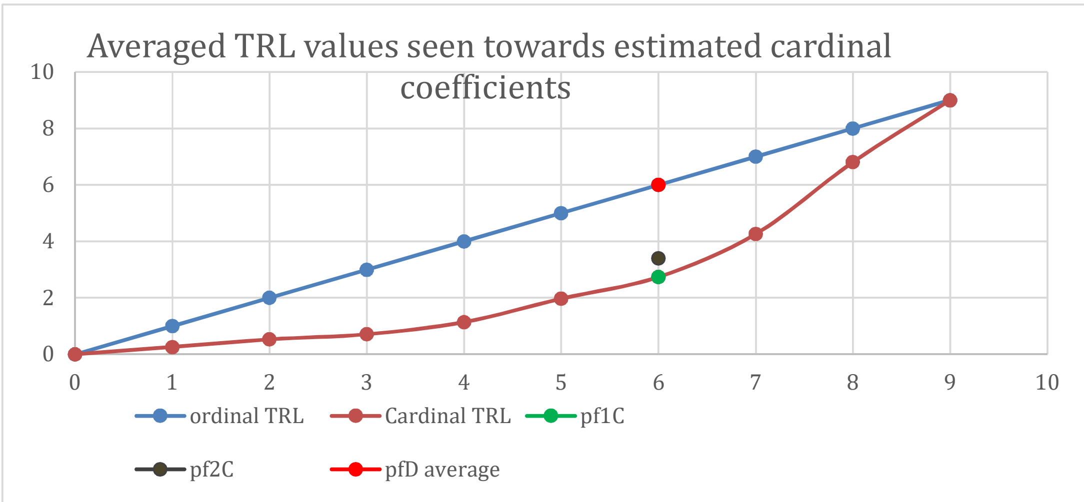

Figure 1 Graph indicating the equivalent cardinal TRL values based on AHP (Conrow, 2009) related to the averaged TRL of two arrays described in text.

Figure 1 indicates that the TRL-heterogeneous project has an average higher degree of maturity when corrected for ordinal scale, but still has a higher likelihood of schedule slippage. This result either underlines the necessity of ensuring all TRLs are as close to TRL 7 as possible, or confirms that the ordinal and subjective characteristic of the scale makes the TRL and other similar scales unsuited as cornerstones in mathematical evaluation giving results that are "computationally accurate, but irrelevant" (Kujawski, 2013).

Steele (2014) published a thesis covering a scalable systems engineering ontology for small and medium sized research organizations with activities in accordance with the severity of the risks. He found that the TRLs can clearly be used to indicate the right focus of systems engineering in order to limit the risks that are found to be linked to the different TRLs.

TRL in a resource perspective. The studies described above, to a large degree, show that the risks indicated by the TRL level of a CTE will affect the performance of projects and programs. Dubos, Saleh, et al., (2008) have considered the additional cost overrun results from early integration, and have proposed that programs should have a margin equal to the mean schedule slippage of a given TRL, since the delays can be anticipated upon early introduction into a program. As an example, the introduction of TRL 6 CTE technology into a system, a 30% margin should be adopted, referring to the Schedule Risk Curve in Figure 2. Specific industries may need to prepare equivalent graphs to be able to calculate the risk margin necessary to facilitate a high-risk project in their own industry.

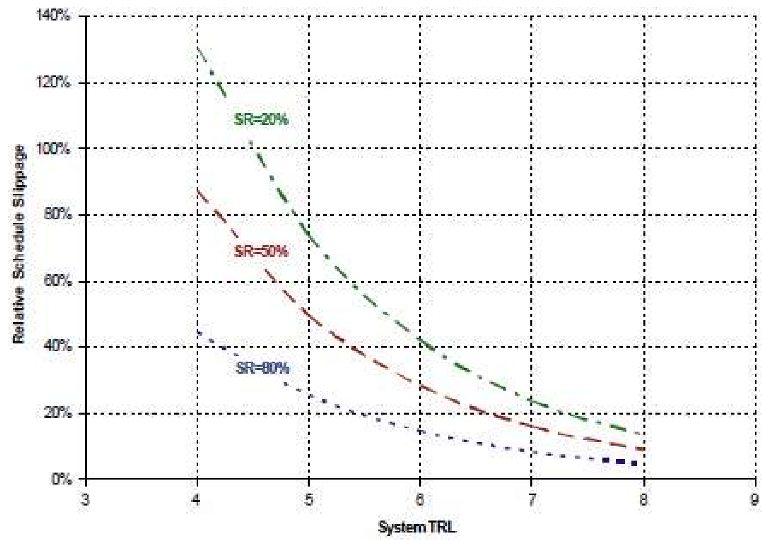

Figure 2 TRL Schedule risk curves (SR) for a normally distributed Relative Schedule Slippage defines the necessary margin to reduce schedule slippage risk (Dubos, Saleh, & Braun, 2008), based on numbers from (Lee & Thomas, 2001)

A weakness of the TRL is the difficulty of defining the amount of resources required to achieve one increment of TRL transition. The expected cost of the next TRL transition should be reassessed on a periodical basis in order to evaluate the technologies likely applicability into systems that will be developed (Gatian, 2015). The recommended acceptance criteria from GAO, TRL 7, was found proven with an average 4,8% cost overrun with all technology matured, while immature technology provided an average 34,9 % cost overrun (Meier, 2008; Katz, et al., 2015). It is apparent that TRL 7 is highly preferable at integration, which further indicates that the decision makers tend to underestimate the cost of maturing technology from TRL 6 to 7. The relative schedule change (RSC) is clearly correlated to the relation between critical design review (CDR) to milestone C (MS C) in the US DoD development process, indicating that a high Design Maturity at CDR is more relevant than the TRL, (Katz, et al., 2015).

By defining the CTE for the project or program according to the TRL scale, the maturity of an individual element may indicate possible future failure of the mission unless a sufficient degree of funding or time is added (Dubos, et al., 2008). The TRL provides decision makers with information to take a well-founded decision on whether to stay with a low TRL-level technology with added funding and time, change to a more mature technology, or cancel the program or project. (Gatian, 2015). The TRL scale can provide value by pointing out the total risk reduction in the program, in the cases of early definition of critical components. To conclude the paragraph of TRLs effect on the projects, a reference to early integration such as with state-of-the-art hardware, with a maturity of less than TRL 4, which is validated in parallel with the program, is correlated to compromised mission success (Meier, 2008).

Suitability of use. As mentioned, the use of TRL is extensive throughout organizations such as: NATO (NATO, 2008), EU (EU, 2014), the US Department of Defense (U.S. DoD, 2010), the US Department of Energy (U.S. DOE, 2010), and US Homeland Security, subsea applications through API standard for subsea (Yasseri, 2016; Yasseri, 2013), and certification bodies (DNV GL, 2017), commercial companies such as Boeing (Whelan, 2008), Google, Raytheon, John Deere, Alstom and BP (Olechowski, et al., 2015). Challenges using TRL in different industries have to some extent been considered. The most relevant issues outside of the native domain is the lack of suitable ways to consider system readiness, among others, by the use of considering the lowest TRL as the weakest link. Moving a specific item from TRL 3 to TRL 5 may only require a few simple tests. The 'weakest link method' does not reflect the remaining work or risk as it does not include the degree of difficulty. BP has attempted to handle this by producing a TRL excel chart that also take degree of difficulty into account by color-coding (Olechowski, et al., 2015).

In 2009, a survey was completed where technology developers and managers were presented with the same set of data. The results showed three levels of difference in rating partially based on perspectives – developers looked at the TRL scale from the documentation to provide proof of suitability. The "relevant environment" was pointed out as incongruent, as the developers tended to think that testing in real environments as too costly while program managers considered simulations to be too poor a representation. (Robinson, et al., 2009).

Furthermore a lack of alignment of TRLs with decision gateways has been mentioned when moving out of the Aerospace and Defense domain. Customizations as in the API N17 (Yasseri, 2013), that uses 7 levels, rather than 9 is an alternative. This however also reflects gaps in research; i.e., will the value that can be interpreted from TRL studies performed via NASA and DoD be valid when the scale changes? Google also indicated that the TRL is not suited for their operations and need to map the technology into future product roadmap; however, the lack of a better alternative maintains the continued use of TRLs (Olechowski, et al., 2015). Upon achieving TRL 9_{NASA}, the equipment is expected to be military grade, this indicates that the product should have passed the infant mortality part of the bathtub curve. When turning to consumer markets, the risk of a few failed product units may be unimportant relative to the risk transferred to stakeholders upon ensuring that equipment achieves TRL 9. Pushing the products up the TRL ladder would increase risk of obsolescence. This indicated that the levels of the TRL are not in line with other domains where the main business driver is not quality, but time or cost, where needs may go directly from an acceptable product (in accordance with defined requirements) to a need for a well-managed life-cycle, where failed products can be replaced in a cost efficient manner.

The commercial example indicates a need to customize the TRLs, which is happening in practice. A consequence of the use of customized TRL values is complications for technology transfer or similar exchanges between two entities cooperating cross-domain. Both entities may have a successful implementation of their own version of TRLs. The understanding of the TRL in the sociocultural production environment is not mentioned. The different applications are rarely shown with a denominator identifying "original system", such as TRL_{NASA}, or TRL_{API}, but this could help avoid confusion. A relevant comparison is the use of SI and Imperial units, and reference to previous projects where this has caused confusion such as for the mars orbiter.

Tools such as the TRL calculator (Nolte, et al., 2003) have been published, in an attempt to make TRLs more available and homogenous. A short list of limitations when applying TRL follows.

- *Integration and system views* are not included in the scope of the TRL; it is not possible to define whether interfacing elements are ready for integration to a top-level system. The system functions have not been considered, as they represent more than the sum of the parts, the simplified breakdown into individual technologies does not allow this evaluation to be done, see (Mankins J. C., 2002; Sauser B. , Ramirez-Marquez, Verma, & Gove, 2006).
- *Capability,* the system's ability to produce an operational outcome, is not considered in the TRLs (Tetlay & John, 2009).
- *Individuality of elements and their properties* Non-system technologies, such as processes, methods, algorithms, or architecture are not adequately considered (Graettinger, Garcia, Schenk, & Van Syckle, 2002).
- *Complexity and uncertainty* of an intelligent systems and their functions and assessments; measures cannot be assessed or controlled with TRLs (Meystel, et al., 2003).
- *Risk and degree of difficulty* The TRLs do not assess the risk associated with transitioning to the next level (Mankins J. C., 2009).
- *Continuously evolving systems* will not benefit from the traditional TRLs, (Smith, 2005).

- *Lifecycle aspects* are not considered in the transition to operation; TRL_{NASA} 9, does not give any further information about the systems expected life cycle (Straub, 2015).
- *Unorganized expansion of TRL***s** to adapt the system for consumer market R&D processes has been attempted (Straub, 2015; Hicks, Larsson, Culley, & Larsson, 2009). Some, such as the EU has also included TRL 0 (Schild, 2014).
- *Customization of TRL* - The use of TRLs in API N17 has been translated to follow the process of the governing bodies and development process (Yasseri, 2016). By harmonization with regulatory structures, the TRL scale is likely to provide value for all stakeholders involved.
- *TRL and development models* often do not align. Stages gates (Cooper, 1996) have been proposed to contain or align with TRL transitions, in other cases the spiral model (Boehm, 1988) is applied, where each iteration represents a new TRL. TRLs in combination with a rigid application of development processes have been identified to have a negative effect on high uncertainty projects. (Högman & Johannessen, 2010).
- *Obsolescence* lack of functionality relative to change of expectations as time passes. If a program progresses over decades, technology development will push a change in the requirements and technologies applied. The TRL has no way to indicate likeliness of obsolescence or the need for changes (Valerdi & Kohl, 2004).
- *Subjective judgement* – there has originally not been "one right way" to conclude the TRA, indicating that the subjective opinion of the person doing the assessment may affect the outcome. Within this conclusion lies a person's nominal understanding of the world and their inherent acceptance for risks, which may also affect the result. (Olechowski, et al., 2015; Sarfaraz, Sauser, & Bauer, 2012).
- *Proliferation* of readiness parameters (SRL, IRL, Logistics readiness level, etc.) have emerged due to omissions of consideration for complex systems, value chains, varying frameworks, human factors and other important metrics (Nolte, 2011).

The limitations of the TRL scope has caused a proliferation of alternative readiness levels, some of which are: Manufacturing Readiness Levels (MRL) (Morgan, 2007), Integration Readiness Level (IRL) (Sauser, et al., 2009), Systems Readiness Level (SRL) (Sauser B. , et al., 2006), Design readiness level, Capability readiness level (Tetlay, et al., 2009), Software Readiness level, Human readiness level, Logistics readiness level, Operational readiness level, Innovation readiness level (Lee, et al., 2011) and Programmatic readiness level as methods most commonly used according to U.S. Department of Energy (DOE) (Fernandez, J. A., Sandia National Laboratories, 2010).

Likewise, many qualitative assessments such as SRL (Sauser et al., 2006), TRRA, ITAM (Mankins, 2002), TRL for non-developmental item (NDI) Software (Smith, 2005), TRL Schedule Risk Curve (Dubos et al., 2008) are based on a TRL evaluation, therefore the relative importance of TRL increases as it has become a cornerstone for other evaluation systems. There exist other methods not based on TRL; yet the TRL scale remains an important scale due to the widespread use in governmental and multinational commercial organizations.

Content analysis

Throughout the documents describing the advantages and disadvantages of applying TRLs in new product development processes, the TRLs are often cited and described. In order to gain an understanding of what was the "right" TRLs or what degree of customization was applied an overview of the TRLs contents and descriptions was created. Table 1 is a sample of the summary tables generated. The subsequent survey phrases were chosen based on the different TRL definitions found throughout academic and user organization literature. When looking through these tables, it appears that TRL 5 (shown in Table 1) is the most diverse when described. TRL 6 that in many cases is just considered a more complicated version of TRL 5, is the least diverse.

Table 1 results from evaluation of description of TRL 5 in different documents gathered from research papers and user organizations, where words are mentioned in at least 4 sources

The references used in the table below are shortened (coauthors left out) due to space restrictions. Similar tables were generated for each of the other TRL documented in 4 or more sources.

descriptive words	Mankins 1995	Mankins 2002	Meystel 2003	Sauser 2006	Hicks 2009	Mankins 2009	Straub 2015	Nolte cal	NASA 2002	NASA 2012	GOA, 1999	DoD 2005	DoE 2008	DOE 2011	DOE CCSI 2012	NATO 2006	EU Horiz. 2014	Boeing 2008
validat*	X	X	X	X	X	X	X	X	X		X	X		X	X	X	X	X
Component	X	X	X	X		X	X	X	X	X	X	X		X	X	X		X
relevant environment	X	X	X	X	X	X			X	X	X	X		X	X	X	X	X
test	X	X			X	X	X	X		X	X			X	X			
breadboard	X	X	X	X		X		X	X	X	X	X				X		X
Integrat*	X	X				X	X	X			X			X	X			
simulat* environment	X	X				X	X	X		X	X			X	X			
realistic	X	X	X			X	X	X							X			
demonstra*	X	X				X	X	X							X			
design							X	X							X			
interfac*						X	X	X							X			
high (-)fidelity								X			X			X	X			
Process								X					X		X			
component level	X	X				X	X											
subsystem level	X	X				X	X											
system level	X	X				X	X											

Survey Results

A majority of the respondents were handpicked because they work within new product development. Some responses were received from a request placed on the official INCOSE LinkedIn site. The industry domain of the respondents is distributed as follows: Medical 1; Maritime 2; Energy: 6; Oil & gas subsea 4, oil &gas topside 4, Software 6, General industry 5, Defense 6, Other 3. Fifty-six percent of the respondents answered that technical maturity was very relevant for work they do. The respondents are 13 experienced in R&D, engineering or design, while 8 work as project managers and 9 work as technical managers. To some extent these roles are all within the target group that use the TRL; research personnel, research management and program management (Sadin, Povinelli, & Rosen, 1989). The survey was intended to give a fresh viewpoint from project employees, valid for organizations where individual responsibility among the employees is emphasized. The TRL documentation available is to a high degree based on defense and aerospace organizations, where the rank and responsibility is clearly defined, relative to many smaller companies, where people are less likely to be specialized and may share a larger portion of the responsibility.

Knowledge of TRL: The respondents in the survey have indicated their knowledge about TRL levels. The median value is 2, while the average is 2,7. The scale ranged from 1 to 5, where 1 indicates 'no knowledge' and 5 is 'experienced user'.

Analysis: The survey was composed of 18 statements from the TRL tables that were typical descriptors for TRL levels. The TRLs 1 and 9 are straightforward and received the greatest number (57%) of correct responses. After crossing TRL 4, the distribution is less focused, indicating some confusion concerning the content of each individual TRL level. The most relevant part of this result is the variation around TRL 5-7. If a discussion is ongoing within a program, and a number of people are participating – it is unlikely that the participants are able to individually identify what a TRL 5, TRL 6, or TRL 7 is without further specification or discussions. It appears that it is around the levels TRL 5, TRL 6 and TRL 7 that statements are the most likely to be misinterpreted. This is in consistent with the increasing importance of the individual TRL levels, as mentioned in the literature review.

It is evident from the questions shown in Figure 3 that the respondents think that technology maturity is a useful measure; however, this raises questions with regards to why it is not all that commonly used in general industry, as stated in question 4.

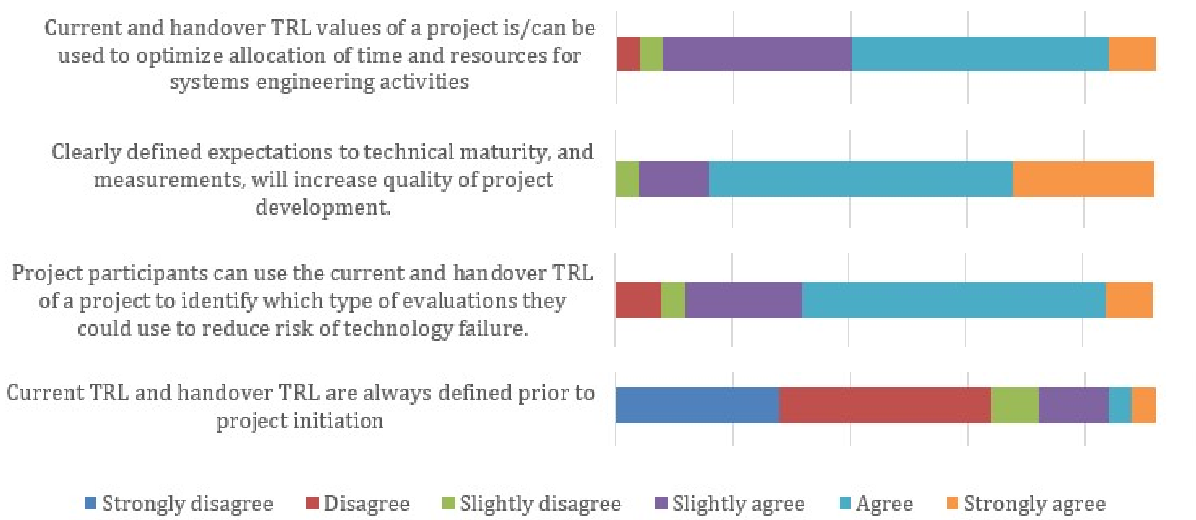

Figure 3 Question from questionnaire on a rating from 1-6

Discussion

The discussion begins with an evaluation of the results as they address the research questions.

TRL can be stated in requirements, does this drive the need for systems engineering?

When the TRL for the end product is stated as a requirement from stakeholders it generates a need to validate that the level has been reached; an expectation of the functional level of the end product. Meier (2008) specifies that investment in Systems Engineering is vital to avoid future cost slippages; and concludes that: "Systems engineering serves as the glue that binds the technical solution to the high-level requirements and maintains the program baseline." The results of his study show that "inadequate systems engineering, coupled with a broken requirements process and immature technology, inevitably leads to program failure" (Meier, 2008). If TRL is stated as a requirement, the needed level for verification or validation of this TRL should be reflected in the wording of the other high level requirements. This indicates a need for thoroughly defined requirements to ensure that the TRL requirement and the functional requirements are aligned and do not contradict or over specify relative to the achievements required in the TRL. Systems engineering becomes relevant in the cross section where complexity meets a need for reliable plans and outcome quality (Honour, 2011). The readiness level transitions in a project could give an indication of additional systems engineering tools to use and the need to establish a technical performance monitoring plan.

TRL can be used to determine risk in the program, does this drive the need for systems engineering?

Every project begins with a set of known and unknown uncertainties, indicative of an initial set of risks. The risks that are not defined and mitigated or otherwise handled become a part of the fabric of the project under the 'cone of uncertainty' (Boehm, 1988). Identifying TRLs will not change the initial risks of the project. However, project management can use the TRL as a part of the feedback loop; when assessing whether the project is on the right track. This evaluation may be used for decision-making where a higher degree of systems engineering may be required to limit uncertainty or reduce risks. If the project is following the expected route and the risk is considered under control, no additional incentives will be necessary. Therefore, the TRL cannot be seen as a driver for systems engineering, however it can be seen as a way of pointing out the need to make adjustments, if the overall risk for CTEs appears to be unacceptable.

The respondents were asked about the relation between risk mitigation, systems engineering and TRLs; the results show that there is in fact a high degree of confidence that maturity metrics can provide valuable information on how to handle project risks. Systems engineering offers a "technical framework for conducting trades among systems performance, risk, cost and schedule" (Reed, 2009). Therefore, TRLs can be seen as metric measurement in the feedback loop that guides the systems engineering process, and by extension, value engineering is implemented into the systems engineering process. Value engineering is a field of engineering that systematically evaluates functions of systems, equipment, facilities, services and supplies for the purpose of achieving the lowest lifecycle cost consistent with required performance reliability, quality and safety (Mandelbaum, et al., 2006). The intention should be to provide well defined quantitative check points for management to respond in a short time, as a way integrating business objectives into the end to end product development processes (Holmes, et al., 2004).

TRL often contributes to determination of budget and schedule parameter, does this drive the need for systems engineering?

To be able to consider the value of SE in specific projects, it is necessary to consider that not all projects have the same end. At a target company, some projects deliver documentation equivalent to a technical maturity TRL 2 while other projects deliver TRL 6 or TRL 9 systems. Logically, the percentage of resources invested in SE should differ in each case, as an early phase project is likely to need a considerably higher percentage of SE than a later phase project under ideal conditions. The use of TRL has also been used to indicate the real value of possible investments (Shishko, et al., 2004). Regardless of TRL, the value of systems engineering has been evaluated by comparing the percentage of the resources applied towards the end cost and schedule for projects. The conclusion shows that systems engineering decreases cost and schedule overruns. Honour (2011) determined that the optimum SE effort appears to be 15-20 percent of the total project cost. However, he also notes that the data he collected showed normal budgets to be only 3-8 percent. The value of Systems engineering is therefore considered proven, however the correlation between value of systems engineering and the TRL level at which the products are delivered is unexplored, and it may vary greatly with the project type and degree of difficulty.

Meier (2008) has conducted research on behalf of different US government agencies and concludes with a need to invest in systems engineering. Meier (2008) specifies requirement creep to be the largest single factor to effect end results, even if technical maturity is also considered a contributing factor. Some of the problems of pointing towards the need for investment in systems engineering is the expansive definition of the discipline itself, the systems engineer performs several roles in a project, contributing with project management tools, through requirements ownership, systems design, systems analysis, validation and verification, logistics, glue among subsystems, interface, technical managers, etc. (Sheard, 1996). It is possible to overspend in one area, while neglecting another. Systems engineering is a technical discipline, however differing from electrical and mechanical and other engineering disciplines; the lack of a physical end product makes the value of systems engineering difficult to quantify (Valerdi R. , 2005). Based on Meier's conclusion there is an apparent need to increase general spending on requirements ownership, however the other roles are not specifically identified. Yet with a broken requirement, it may be difficult to design the right system and validate the correct system function. This indicates a substantial degree of dependence between the different systems engineering roles. In turn, the effect of systems engineering spending will have a cascading effect from the start of the project. The correct spending at the right TRL level as an indicator could help by ensuring that the start of the project is handled sufficiently well; ensuring a good basis for the other systems engineering roles to be able to perform. It could be that this kind of evaluations could be related to the Steele's scalable ontology (Steele, 2014), where combining project details in a database could give an indication of optimal spending within the organization.

How does this information affect other industries?

In automotive, and to some extent industry in general, the lean approach is more common than in typical systems engineering domains (Welo, et al., 2015). When responsibility for improvements trickles down to the employees, as is normal in lean companies, the need to understand the global view increases.

It is possible to consistently understand and apply a maturity metric. Checking a single digit number as an indication of maturity is easy enough for any project employee, sales representative or other role. Employees may be motivated if they are aware that they may be saving the company an average 30% cost overrun (Meier, 2008). The functions of the TRLs are to ensure that a communicated level of maturity has been reached. A successful level of communication around TRL requires that all parties have the same understanding of what a TRL level entails.

Recommendations for future work

Maturity metric standardization - The TRL is already defined in a set manner. The TRLs likely serve as a very good beta-test for a tool that is suitable for all domains with varying degrees of complexity. It may be that the best evaluation method can be found one meta-level up, and defines a set of evaluations that is applicable to different systems, times, places, organizations and that would give the relative end level that was relevant to any specific product. The ideal solution would be a modular evaluation system, where the different layers were chosen based on the specification of the product connected to business drivers.

Development - A TRL level in its semantic cloak is not necessarily the same today, as it was in 1995. By gathering more user organizations data; to query about what they look for; and compare it towards the definitions from literature would be a start to improve the current understanding. There is a continuous development both of work methods, tools, and technology, and society. The past 10 years there has been an emergence of platform structures in society, where rather than a top down organization, people work in flat organizations. The tools and structures humans work in correlate to the way work is done. When society changes, so must the ways people work change in order to stay relevant. The TRL scale is not omitted from this reality.

The human aspect in an organization - Under the technical literature that is overarching the TRL concept, there appears to be one wall missing. The regional culture, as well as organizational culture may affect the application of the scale.

Conclusion

Based on the results presented in Table 1, the understanding of each individual technology readiness level varies, and is not conclusive without a context and a commonly agreed viewpoint. The TRLs are not intuitive, and prohibit clear communication regarding technology maturity when they are presented without reasoning, discussion, or context. It may be useful to go through the definitions in advance of discussions of TRL levels. An alternative is to create a visualization, such as a poster, that defines the content for the personnel to which it is relevant, as a way of internalizing the scale.

Specification of reference scale so that context is always maintained could be recommended, exemplified by e.g. $TRL7_{NATO}$ and $TRL7_{API\ 17N}$. The TRLs are not guaranteed to be equivalent to each other and this customization may ultimately cause confusion. Transfer of technology is stated to be a highly-prioritized activity among both NASA and ESA, lacking this specificity may cause problems in complex cross domain projects.

Besides the possibility of misunderstandings TRLs can provide a frame to help identify suitable systems engineering tools after identifying the risk profile of the project. The application of TRLs in lean organizations may require other strategies than what TRLs were originally intended to cover. The TRL was created for management and R&D to communicate. In a lean organization, where responsibility is percolating down through the organization the need for the employee to see the bigger picture is greater. The use of TRLs requires a lot of knowledge about the TRA and indicates that it may be a less suited metric in the lean organizations. It is likely that the TRA works better in more hierarchical organizations, where specialization means that not as many persons are involved in the TRA.

References

Austin, M. F., Homberger, C., Polacek, G. A., Doolittle, E., Ahalt, V., & York, D. M. (2018). Using Bayesian Networks to Validate Technology Readiness Assessments of Systems. In Disciplinary Convergence in Systems Engineering Research (pp. 787-797). Springer, Cham.

Azizian, N., Sarkani, S., & Mazzuchi, T. (2009). A comprehensive review and analysis of maturity assessment approaches for improved decision support to achieve efficient defense acquisition. Proceedings of the World Congress on Engineering and Computer Science 2009 Vol II. San Francisco.

Boehm, B. W. (1988). A spiral model of software development and enhancement. IEEE Computer, 21(5), 61-72.

Conrow, E. H. (2009). Estimating technology readiness level coefficients. AIAA Paper No. 6727.

Cooper, R. G. (1996). Overhauling the new product process. Industrial Marketing Management, 25, 465-482.

Dacus, C. (2012, October Vol. 19, No. 4). Improving Acquisition Outcomes Through Simple System Technology Readiness Metrics. Defense ARJ, pp. 444-461.

Davis, D. e. (2015). Performance of Defense Acquisition systems. Washington, https://www.defense.gov/Portals/1/Documents/pubs/Performance-of-Defense-Acquisition-System-2015.pdf: U.S. Department of Defense.

DNV GL AS. (2017, 05 04). www.dnvgl.com. Retrieved from DNV-RP-A203: https://rules.dnvgl.com/docs/pdf/DNV/codes/docs/2011-07/RP-A203.pdf

Dubos, G. F., Saleh, J. H., & Braun, R. (2008). Technology readiness level, schedule risk, and slippage in spacecraft design. Journal of Spacecraft and Rockets, 45(4), 836-842.

Dubos, G., & Saleh, J. H. (2011). Spacecraft technology portfolio: Probabilistic modeling and implications for responsiveness and schedule slippage. Acta Astronautica, 68, 1126–1146.

Edmondson, A. C., & McManus, S. E. (2007). Methodological Fit in Management Field Research. Academy of Management Review, 32(4), 1155-1179.

Engel, D., Dalton, A., Anderson, K., Sivaramakrishnan, C., & Lansing, C. (2012). Development of Technology Readiness Level (TRL) Metrics and Risk Measures. Richland, Washington: Pacific Northwest National Laboratory, Prepared for US. Department of Energy .

ESA, TEC-SHS. (2008). Technology Readiness Levels handbook for space applications, issue 1, rev. 6. Retrieved from https://artes.esa.int/sites/default/files/TRL_Handbook.pdf

EU. (2014). ec.europa.eu HORIZON 2020. Retrieved 05 04, 2017, from HORIZON 2020 – WORK PROGRAMME 2014-2015, G. Technology readiness levels (TRL)
Fernandez, J. A., Sandia National Laboratories. (2010). Contextual Role of TRLs and MRLs in Technology Management. Albuquerque: Sandia National Laboratories operated for the United States Department of Energy.
Gatian, K. N. (2015). A Quantitative, Model-Driven Approach to Technology Selection and Development through Epistemic Uncertainty Reduction. Ph.D thesis, Georgia Institute of Technology, School of Aerospace Engineering.
Graettinger, C. P., Garcia, S. S., Schenk, R. J., & Van Syckle, P. J. (2002). Using the Technology readiness scale to support technology management in the DoD ATD/STO environments. A finding and recommendation report. CMU/SEI-2002-SR-027. Pittsburgh, PA 15213-3890: Carnegie Mellom Software Engineering Institute.
Hicks, B., Larsson, [., Culley, S., & Larsson, [. (2009). A Methodology for Evaluating Technology Readiness During Product Development. International Conference on Engineering Design. Stanford University, Stanford, CA, USA:
Holmes, M., & Campbell, R. (2004). Product development processes: Three vectors of improvement. Research Technology Management, 4, 47-56.
Honour, E., (2011). Sizing Systems Engineering Activities to Optimize Return on Investment, INCOSE International Symposium, Denver, CO, USA, 2011.
Högman, U., & Johannessen, H. (2010). Technology development and normative process models. International Design Conference, (pp. 265-274).
Katz, D., Sarkani, S., Mazzichi, T., & Conrow, E. H. (2015). The Relationship of Technology and Design Maturity to DoD Weapon System Cost Change and Schedule Change During Engineering and Manufacturing Development. Systems Engineering, Vol. 18 No.1 pp. 1-15.
Kujawski, E. (2013). The trouble with the System Readiness Level (SRL) index for managing the acquisition of defense systems (presentation). IEEE Transactions On Systems, Man, And Cybernetics: Systems, 43(4).
Lee, M. C., Chang, T., & Chien, W. T. (2011, May). An approach for developing concept of innovation readiness levels. International Journal of Managing Information Technology (IJMIT), 3(2), 18-38.
Lee, T. S., & Thomas, L. D. (2001). Cost Growth Models for NASA'S Programs: A Summary. Interfacesymposia.org.
Mandelbaum, J., & Reed, L. D. (2006). Value Engineering handbook. Institute for defense analyses.
Mankins, J. C. (1995). Technology Readiness Levels: A White Paper. Retrieved 4 7, 2017, from NASA, Office of Space Access and Technology, Advanced Concepts Office.
Mankins, J. C. (2002). Approaches to Strategic Reseach and Technology (R&T) Analysis and Road Mapping. Acta Astronautica, 51.
Mankins, J. C. (2009). Technology readiness assessments: A retrospective. Acta Astronautica, 65, 1216-1223.
Meier, S. R. (2008). Best project management and systems engineering practices in the preacqusition phase for federal intelligence and defense agencies. Project Management Journal, 39,(No 1), 59-71.
Meystel, A., Albus, J., Messina, E., & Leedom, D. (2003). Performance Measures for Intelligent Systems: Measures of Technology Readiness1. PERMIS'03 White Paper.
Morgan, J. (2007, 11). Manufacturing Readiness Levels (MRLs) and Manufacturing Readiness Assessments (MRAs). (Air Force Research Lab Wright-patterson Afb Oh Manufacturing Technology Directorate)
NASA. (2012, 10 28). NASA. Retrieved 05 2017, from Technology Readiness Level: https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_accordion1.html
NATO. (2008). A Holistic Approach for NATO Research & Technology.
Nolte, W. (2011). Readiness Level Proliferation. AFRL (Air Force Research Laboratory).
Nolte, W. J., Kennedy, B., & Dziegiel, R. J. (2003). Technology Readiness Calculator.

Olechowski, A., Eppinger, S. D., & Joglekar, N. (2015). Technology Readiness Levels at 40: A Study of State-of-the-Art Use, Challenges, and Opportunities. SSRN Electronic Journal.

Reed, D. M. (2009, May-June). Value Engineering throughout a defense system's life cycle. Defense AT&L.

Robinson, J. W., Levack, D. J., Rhodes, R., & Chen, T. (2009). The need for technology maturity of any advanced capability to achieve better life cycle cost (LCC). 45th ALkA/ ASME/ JSAEJ/ ASEE Joint Propulsion Conference. Denver, Colorado.

Sadin, S. R., Povinelli, F., & Rosen, R. (1989). The NASA technology push towards future space missions. Acta Austronautica Vol. 20, 73-77.

Sarfaraz, M., Sauser, B. J., & Bauer, E. W. (2012). Using system architecture maturity artifacts to improve technology maturity assessment. In C. H. Dagli (Ed.), Procedia Computer Science, 8, pp. 165-170. St. Louis, MO.

Sauser, B. J., Long, M., Forbes, E., & McGrory, S. E. (2009). Defining an Integration Readiness Level for Defense Acquisition. INCOSE Int'l Symposium, Vol 19, Issue 1. Singapore.

Sauser, B., Ramirez-Marquez, J. E., Magnaye, R., & Tan, W. (2008). A systems approach to expanding the technology readiness level within defense acquisition. International Journal of Defense Acquisition Management, 1, pp. 39-58.

Sauser, B., Ramirez-Marquez, J., Verma, D., & Gove, R. (2006). From TRL to SRL: The Concept of Systems Readiness Levels. Paper #126, Stevens Institute of Technology.

Schild, P. (. (2014). HORIZON 2020 Calls - Overview.

Sheard, S. A. (1996). Twelve systems engineering roles. INCOSE sixth annual international symposium. Boston, MA.

Shishko, R., Ebbeler, D. H., & Fox, G. (2004). NASA Technology assessment using real option valuation. Systems Engineering, The journal of the International Council of Systems Engineers, 7(1).

Smith, J. (2005). An alternative technology readiness level for non-developmental item (NDI) software. Proceedings of the 38th Hawaii International Conference on Systems Sciences, January 3-6, 2005, Island of Hawaii, USA.

Steele, R. J. (2014). Scalable Systems Engineering Ontology for Small and Medium Enterprise (SME) Research Organizations, Ph. d. thesis.

Straub, J. (2015). In Search of Technology Readiness Level (TRL) 10. Elsevier, 312-320.

Tan, W., Ramirez-Marquez, J., & Sauser, B. (2011). A probabilistic approach to system maturity. Systems Engineering, 14(3).

Tetlay, A., & John, P. (2009). Determining the lines of system maturity, system readiness and capability readiness in the system development lifecycle. 7th Annual Conference on Systems Engineering Research (CSER 2009).

U.S. DoD. (2005). Department of Defense; Technology readiness assessment (TRA) deskbook.

U.S. DoD. (2009, July). DEPARTMENT OF DEFENSE Technology Readiness Assessment (TRA) Deskbook MEDICAL RDT&E.

U.S. DoD. (2010). www.army.mil. Retrieved 05 04, 2017, from Technology Readiness Levels in the Department of Defense (DoD) - 2010

U.S. DoD. (2011). Department of Defense: Technology readiness assessment (TRA) Guidance.

U.S. DOE. (2010). https://energy.gov. Retrieved 05 04, 2017, from energy.gov - the Department of Energy: https://energy.gov/sites/prod/files/em/Volume_I/O_SRP.pdf

U.S. DoE. (2011). Deparment of Energy: Technology readiness assessment guide.

U.S. DoE. (2011). DOE G 413.3-4A, Technology Readiness Assessment Guide. Retrieved from https://www.directives.doe.gov/directives-documents/400-series/0413.3-EGuide-04a

U.S. GAO. (1999). U.S. Government Accountability Office: Best Practices: Better Management of Technology Development Can Improve Weapon System Outcomes, GAO/NSIAD-99-162.

U.S. GAO. (2009). U.S. Government Accountability Office: Technology readiness assessment guide - Best practices for evaluating readiness of technology for use in acquisition programs and projects.

Valerdi, R. (2005). The constructive systems engineering cost model (COSYSMO). A Dissertation Presented to the Faculty of the Graduate School University of Southern California, 152.
Valerdi, R., & Kohl, R. (2004). An Approach to Technology Risk Management. Engineering Systems Division Symposium. MIT, Cambridge, MA.
Vegvesenet (NPRA). (2016). Risk Management and Technology Qualification, Bridge conference.
Welo, T., & Ringen, G. (2015). Investigating Lean development practices in SE companies: A comparative study between sectors. 2015 Conference on Systems Engineering Research (pp. 234-243). ScienceDirect.com.
Whelan, D. (2008, 08 04). Impact of Technology Readiness Levels on Aerospace R&D (presentation notes).
Yasseri, S. (2013). Subsea system readiness level assessment. Underwater Technology The International Journal of the Society for Underwater.
Yasseri, S. (2016). A Measure of Subsea Systems' Readiness Level. Underwater Technology, The international Journal of the society for Underwater Technology, 115-128.

Biography

Kjersti Bakke has a BSc in Electrical Automation from Vestfold University College in 2006 and an MSc of Systems Engineering from University of Southeast Norway, 2016. She has worked with electrical engineering and instrumentation full time since 2006, through verification work in DNV GL and as a project engineer at Wood plc. She serves as the employee elected representative on the board of directors of Wood Group Norway AS, a subsidiary of Wood plc since 2017.

Cecilia Haskins entered academia after more than thirty years as a practicing systems engineer. Her educational background includes a BSc in Chemistry from Chestnut Hill College, and an MBA from Wharton, University of Pennsylvania. She has been recognized as a Certified Systems Engineering Professional since 2004. After earning her PhD in systems engineering from NTNU, she developed and teaches an overview course in systems engineering with a novel lab. Her research interests include engineering education, and innovative applications of systems engineering to socio-technical problems such as those encountered in software intensive systems and sustainable development.