Technology readiness levels at 40: A study of state-of-the-art use, challenges, and opportunities

MIT Sloan School of Management

MIT Sloan School Working Paper 5127-15

Technology Readiness Levels at 40: a study of state-of-the-art use, challenges, and opportunities

Alison Olechowski, Steven D. Eppinger, Nitin Joglekar

April 1, 2015

This paper appears in the proceedings of
Portland International Conference on Management of Engineering and Technology (PICMET), August 2015.

Technology Readiness Levels at 40: a study of state-of-the-art use, challenges, and opportunities

Alison Olechowski[1], Steven D. Eppinger[2], Nitin Joglekar[3]

[1] Massachusetts Institute of Technology, Dept. of Mechanical Engineering, Cambridge, MA 02139 USA

[2] Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA 02139 USA

[3] Boston University, School of Management, Boston, MA 02215 USA

The technology readiness level (TRL) scale was introduced by NASA in the 1970s as a tool for assessing the maturity of technologies during complex system development. TRL data have been used to make multi-million dollar technology management decisions in programs such as NASA's Mars Curiosity Rover. This scale is now a de facto standard used for technology assessment and oversight in many industries, from power systems to consumer electronics. Low TRLs have been associated with significantly reduced timeliness and increased costs across a portfolio of US Department of Defense programs. However, anecdotal evidence raises concerns about many of the practices related to TRLs. We study TRL implementations based on semi-structured interviews with employees from seven different organizations and examine documentation collected from industry standards and organizational guidelines related to technology development and demonstration. Our findings consist of 15 challenges observed in TRL implementations that fall into three different categories: system complexity, planning and review, and validity of assessment. We explore research opportunities for these challenges and posit that addressing these opportunities, either singly or in groups, could improve decision processes and performance outcomes in complex engineering projects.

I. INTRODUCTION

Whether we consider NASA's future manned mission to Mars or next year's popular electronic devices, cutting edge innovation is achieved in part through the application of new technologies. But with the performance advancements gained from using a new technology comes a great deal of uncertainty and risk regarding the technology's capabilities, limitations, and development trajectory. When technologies are not ready on time, the consequence could be budget overruns, schedule delays, performance shortcomings or even project cancellation. A better understanding of the state of the technology maturity is critical in making good decisions about the injection, development and integration of these technologies in complex engineering projects. The most widely used tool for such maturity assessment is the technology readiness level (TRL) scale.

The TRL scale was first developed at NASA in the 1970s to be a consistent measure of technology maturity. Today the TRL approach is being used in multiple industries and serves broader goals than originally intended. This paper reviews implementation practices and aims to highlight 15 challenges that have been encountered based on evidence collected at seven different organizations. We then discuss potential improvements to industry best practice and identify opportunities for future research.

A. TRL History at NASA

Following initial implementation within NASA in the '70s, the technology readiness concept was first published externally in 1989, as a 7-point scale [1]. In this paper, the TRLs are described as being motivated by "the differing perceptions of the researchers and the mission planners between the intended and actual proof of readiness," with the promise that "a properly planned, thoroughly executed technology research and development program can provide substantive advances at acceptable risk levels." In 1995, NASA published a refined 9-point scale, along with the first detailed descriptions of each level [2].

Shown in Table 1, the scale begins with a technology in a very basic scientific form, and progresses to a technology proven in the operating environment. Thus for a generic technology, the levels describe the demonstration requirement, including environment and technology assembly status, at increasing fidelity to the final operating system. In common use, a technology is assessed on the TRL scale to better understand the progression in readiness of the technology towards eventual operation.

Various definitions of "technology" exist at many different scales; the predominant conceptualization within the TRL-community, at NASA, and used in this paper, is that of a component-technology. The TRL scale is used to assess the maturity of a component (or a principle that will eventually be embodied in a component) that features new materials, scale, or working principles.

Table 1: NASA Technology Readiness Level Scale [3]

TRL	Definition
9	Actual system "flight proven" through successful mission operations
8	Actual system completed and "flight qualified" through test and demonstration (ground or flight)
7	System prototype demonstration in a target/space environment
6	System/subsystem model or prototype demonstration in a relevant environment (ground or space)
5	Component and/or breadboard validation in relevant environment
4	Component and/or breadboard validation in laboratory environment
3	Analytical and experimental critical function and/or characteristic proof-of-concept
2	Technology concept and/or application formulated
1	Basic principles observed and reported

Government and commercial implementations of the TRLs are remarkably similar to NASA's original 9-point embodiment. NASA's most up-to-date TRL documentation is publicly available in their Systems Engineering Handbook, including the TRL scale as shown in Table 1 [3]. Perhaps the most detailed publicly available description of TRL assessment and decision-making comes in a white paper written by a consultant while working at the NASA Space Flight Center [4].

B. TRL Usage Beyond NASA

Starting in 2001, the United States Department of Defense required the use of TRLs in all of its new procurement programs, quickly expanding the adoption of the scale. A variety of industries have now generated customized standard guidelines for using TRLs in complex systems development, including defense, oil and gas, and infrastructure [5]–[7].

Organizations are increasingly mapping TRL goals to their generic system development process, i.e., target TRLs are assigned to some or all gates in a development process. The Department of Defense, for example, has mapped the TRL scale to their System Acquisition Process, as represented in Figure 1 [6]. With such a TRL mapping, expectations of technology maturity are consistent and explicit across projects. This practice also ensures that technology maturity is considered in the decision to pass relevant gate reviews. These reviews are shown as milestones in Figure 1.

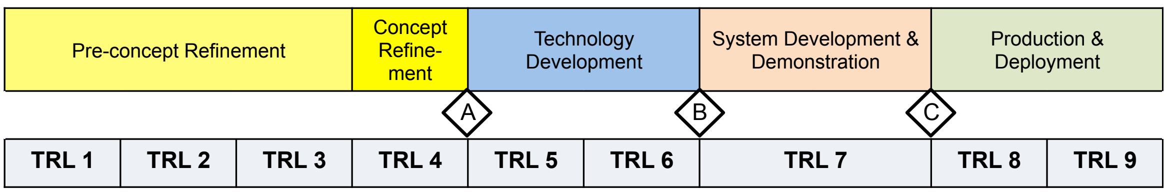

Figure 1: Mapping of technology readiness levels to US Department of Defense System Acquisition Process. Technologies are expected to achieve TRL 4 by milestone A, TRL 6 by milestone B, and TRL 7 by milestone C.

There are a number of reasons why organizations have widely adopted the TRL scale for technology development assessment. One key benefit, as highlighted in Sadin's initial TRL publication, is a shared understanding of technology maturity and risk [1]. The levels are a standard language with which to discuss technology readiness across the organization and between disciplines. They are particularly useful in planning technology hand-offs between different groups, for example a research and development group and a project group. Additionally, the levels provide a systematic approach and model for technology-intensive system development, with the TRLs acting as basic guideposts and steps.

Although limited, there exists some evidence to suggest that the mapping of TRLs to the system development lifecycle is a helpful best practice. A study of 62 US Department of Defense programs found that those programs which reached TRL 7 or higher by the start of system development finished practically on time and on budget, whereas those programs with technologies below a TRL 7 showed, on average, development cost growth of 32%, acquisition unit cost increase of 30%, and schedule delay of

20 months [8]. Another study of 37 Department of Defense weapon systems showed that technology maturity guidance had a statistically significant effect on the schedule overrun of these systems [9]. These findings encourage our further investigation of TRL usage in current practice.

Safety critical industries like defense, oil and gas, and aerospace are increasingly focusing on TRLs as a tool for technology qualification, leading to expanded roles for independent consultants and accreditors, and the writing of new standards. This growth will likely continue, considering that TRL use is increasingly incentivized by grants and funding structures that use the scale as a basis for eligibility, for example the America Makes additive manufacturing initiative, or the European Commission's Horizon 2020 research and innovation programme [10], [11].

C. TRL-Related Literature

Beyond the industry guidelines published on TRL usage, there exists some academic literature exploring technology readiness and related topics. Since his first TRL paper in 1995, Mankins has published further. One paper offers an extended methodology of TRL assessment for effective technology management [12]. Mankins also published a 30-year retrospective on the TRLs [13]. Our work aims to expand the scope of this retrospective, and provide more detailed observations based on evidence collected from diverse industries.

System architecture-related extensions to the TRL include the integration readiness level (IRL) and system readiness level (SRL), introduced by Sauser and his colleagues [14], [15]. This work reflects the reality that technologies do not exist independently, but rather they are connected through interfaces in the system architecture. The IRL uses a 1-9 scale in the style of the TRL to assess the maturity of the interface connecting two components. The SRL is a 0-1 value computed from the system's TRLs and IRLs. The authors mapped SRL values to standard systems engineering life cycles, including those of NASA and US Department of Defense.

Jimenez and Mavris make a case against the use of SRLs based on concerns about the generalizability of the IRL, the assumption of independence of the TRL and IRL, and the mathematical averaging and aggregation required to compute the SRL [16]. They address this criticism by proposing an alternative integration-focused extension to the TRL. This extension comes in the form of integration- and architecture- focused elaborations to the TRL descriptions at each readiness level. The authors argue that integration is an inherent sub-attribute of technology readiness and so the TRL levels themselves should reflect integration readiness.

The visual combination of system architecture information with technology readiness metrics has been explored in two separate works as a tool for system assessment and management decision-making. Both build upon the system architecture design structure matrix [17]. Demonstrated with case studies of NASA's Mars Pathfinder and Near Earth Asteroid Rendezvous, Brady introduced the technology risk design structure matrix [18]. For each of the system's interfaces, this matrix indicates a product of the interface strength and inverted technology maturity for both of the interfacing technologies. A high entry in the matrix could indicate an important interface with high technical uncertainty. In the context of the oil and gas industry, Yasseri presents Sauser's IRL in a system architecture design structure matrix [19]. He then uses the matrix to facilitate the calculation of Sauser's SRL.

Focusing on maturity assessment methodologies, the work of Azizian, Sarkani and Mazzuchi includes a discussion of the strengths and weakness of a variety of maturity-related techniques, such as the IRL, R&D degree of difficulty, and the SRL [20]. Related work on technology readiness assessments (TRA) from the same authors explored the relationship between the act of conducting a formal TRA and system quality or program performance, but found no significant correlation [21].

Cornfield and Sarsfield argue that current TRL measurement techniques are highly qualitative, and the importance of language and culture are generally greatly underestimated [22]. Therefore the authors present a new index linked to engineering requirements that aims to increase the accuracy of readiness estimations.

Our work has been informed by the careful thought found in the literature cited above. We considered the shortcomings of TRL practice raised in these works, and reflected on the motivation for the various TRL extensions. Our study aims to provide empirical evidence for these, as well as previously unpublished TRL-related challenges. In this paper we share the findings of a multi-industry study of state-of-the-art use of TRLs. We present our findings as a set of 15 challenges derived from interviews and documentation review.

II. METHOD

Our study aims to identify and describe the most important challenges relating to use of TRLs in practice today. According to the methodological fit archetypes of Edmonson and McManus, this research fits the “nascent” archetype for field research, given the state of prior research on complex systems development and managers’ use of TRLs in industry [23]. The appropriateness of this choice is further confirmed by the open-ended nature of our research questions. We follow the recommended methodology where the type of data collected is qualitative, with a goal of pattern identification based on thematic content analysis of these data.

A. *Collection of Evidence and Data Analysis*

Evidence was collected from seven technology development organizations. The set of organizations in our study was selected with the aim of industry diversity. The variety of industries results in evidence based on processes of various development lifecycle lengths, regulatory oversight, and competitive environments. Further, the organizations studied reflected a spectrum of degree of maturity of their technology readiness processes, although all individuals interviewed had experience using TRLs. For example, one organization has only just considered formal readiness assessment within the past year, while NASA has been using the scale and refining and expanding its processes since the TRLs were developed there in the 1970s.

Two main sources of data were pursued: (1) interviews, and (2) company- and industry-specific technology development guideline documents.

Descriptive details of the interview participants are provided in Table 2. In all cases, we interacted with those parts of the organization concerned with hardware development; for instance, in the case of

Google, our contacts were not software developers but rather engineering managers working on consumer electronics in the Google X division.

Table 2: Interview participants

#	Organization	Industry	TRL Experience	Role of Interviewee
1	NASA	Space	> 8 years	Office of the Chief Engineer Office of the Chief Technologist
2	Raytheon	Defense	> 8 years	Director of Engineering
3	BP	Oil & Gas	2 – 8 years	Technology Leader Engineering Manager Technology Manager Engineering Manager VP Technology Project Manager Independent Consultant
4	Bombardier	Aircraft	2 – 8 years	Senior Engineering Specialist Systems Integration, Advanced Design
5	John Deere	Heavy Equipment	2 – 8 years	Systems Engineer Systems Engineering Manager
6	Alstom	Power Systems	< 2 years	System Engineer Risk Expert Process Expert
7	Google	Electronics	< 2 years	Program Manager Product Design Lead

Within each organization, we conducted semi-structured interviews with those individuals directly responsible for and involved in technology maturity assessment. In four cases we were able to speak with managers responsible for setting TRL usage guidelines at their organization. Whenever possible we interviewed employees from a variety of roles (project manager, systems designer, technical executive) within the organization to gain a complete perspective of the challenges. Additionally, when possible, we obtained documentation on formal procedures for technology readiness assessments or technology management decisions.

The evidence was coded and each indication of a TRL-related challenge was identified. The evidence was then grouped according to similarity of concept. We considered several different groupings, and ultimately chose a set of 15 groups of evidence that were thought to be most internally consistent and conceptually distinct. Several of the challenges that emerged are reflected in the previously discussed TRL-related literature.

III. FINDINGS: CHALLENGES AND OPPORTUNITIES

15 challenges of modern TRL implementation were identified from the collected evidence. For clarity of presentation, we grouped the challenges into three categories, as shown in Table 3: system complexity, planning and review, and assessment validity. In several cases where a challenge relates to more than one category, we have placed it according to our strongest evidence. In the next section we provide a brief description of each challenge and a representative sample of the supporting evidence.

Table 3: Challenges encountered in modern TRL implementation

System Complexity
1. Integration and connectivity 2. Interface maturity 3. Scope of the TRL assessment 4. Influence of new components or environment 5. Prioritization of technology development efforts 6. System readiness 7. Visualization
Planning and Review
8. Aligning TRLs with system development gates 9. Waivers 10. Back-up plans 11. Effort to progress 12. Confidence to progress 13. Product roadmapping
Assessment Validity
14. Subjectivity of the assessment 15. Imprecision of the scale

A. System Complexity

The following challenges are those that relate to the complexity of the system under development, whether it be the component technologies, the architecture, or the system as a whole.

Challenge 1 - Integration and connectivity

Description: The TRL scale is designed to assess each component technology independently; however in reality, the components are integrated to work as a complete system. During development, demonstration escalates from evaluating an isolated component, to many components in a subsystem, and later the full system. Since components work together to achieve system functions, a component's specification and design are likely to be coupled to those of its neighbors, meaning that changes in one component may require changes in connected components. Integration issues are a key cause of delay and budget overrun yet the TRL measurement does not give us specific integration guidance.

Evidence: A systems engineer at Alstom described why he would like the TRL assessment to better consider integration: "As far as we're concerned, the technology simply is not at a suitable readiness unless it integrates with what's around it." However they have not yet been able to formalize this integration concept in their TRL procedures. An engineer at Bombardier explained that "even though we didn't necessarily have a problem with using TRL in a classic NASA way, we believe – whether it's true or not – that commercial jet aircraft and business aircraft are particularly integral products. So many things affect other things and so integration is a big deal."

Challenge 2 - Interface maturity

Description: Complex systems can be viewed as connected components, where each pair of connected components is joined via an interface. TRLs assess the components themselves but do not explicitly evaluate the maturity of the interfaces. Operational failures are known to sometimes occur at the interface, whether that interface is physical or signal. A challenge exists in effectively considering interface maturity in the technology readiness assessment process.

Evidence: Interfaces between components are often the site of development issues. As a technology manager at BP put it, "In my experience, having worked in the industry for 18 years, a lot of the big mistakes or problems I've seen have been due to poor interface management." Yet a consultant in the oil & gas industry described how in his experience TRL assessments are typically performed independently for each component without regard for interfaces. He explained that "it is not unusual that two pieces of equipment mismatch or require several iterations to work together." Despite their importance, consideration of the interfaces is lacking in technology readiness assessments.

Challenge 3 - Scope of the TRL assessment

Description: Two major scoping approaches prevail in TRL assessment in industry: 1-The Product Breakdown Structure (PBS) approach considers the complete system but requires review of each component, whether risky or not, and is therefore resource intensive. 2- The Critical Technology Element (CTE) approach is more efficient in review and information generated since it considers only components "that may pose major technological risk during development" [6]. However, there exists the chance that a risky technology is left out of the assessment. When establishing a process for TRL assessment, it is unclear how to decide which approach to take, and how this decision should be informed by the complexity of the system, the cost pressures of the project, or the risk of operation.

Evidence: The PBS approach has been adopted in the oil & gas industry, and comes with some downside, as a consultant explained, "operators use TRL not to manage key technologies, but for tracking readiness of all equipment for installation. Every nut and bolt of every equipment is included in an Excel sheet. You can imagine such a spreadsheet will become very large."

At Raytheon, the CTEs are identified prior to the first technology readiness assessment, but less than expected efficiency is achieved since the list does not necessarily remain static, as we learned from a director: "if the system architecture has changed in a manner that introduces new subsystems or new implementation methods, then the team is required to look into those areas and determine if there are new CTEs."

For both approaches, the benefits and weaknesses of the manner in which scope is established remain unclear.

Challenge 4 - Influence of new components or environment

Description: Most TRL guidance is written in the context of a new system development project, where each component technology is assessed at the start and developed over the system lifecycle. In practice, systems are often incremental improvements from previous versions, with a proven system being adjusted slightly to suit a new environment, or where one new component replaces an old component in a heritage system for added performance capability. Organizations struggle with the appropriateness and procedure of the TRL assessment on such development projects.

Evidence: Both NASA's Systems Engineering Handbook and Bilbro's white paper agree that a component can only be TRL 9 (the highest level) if it has been demonstrated in the identical configuration in an identical environment [3], [4]. The NASA guidebook even includes a flowchart (part of which is shown in Figure 2) to elaborate on this point. In our interview with NASA, however, we learned that this challenge is not as resolved as it appears in the guidance, as we heard that "An item we're considering is the effect of hierarchy to TRL – that is, if you have a TRL 9 vehicle but you replace something at the assembly, component or even piece part level with a lower TRL, does the entire system then become that lower TRL?" The NASA guidance suggests the answer is yes, the whole system becomes a lower TRL, but this is clearly not a commonly agreed upon conclusion.

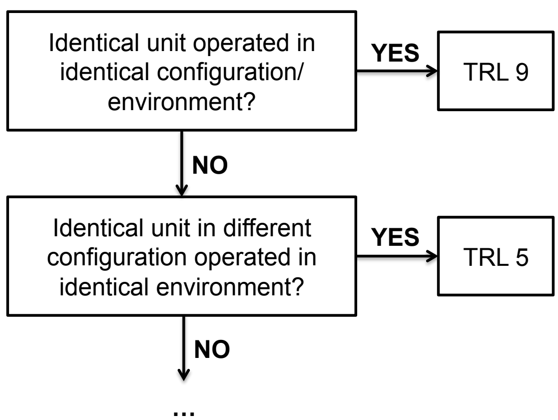

Figure 2: Top section of the decision flowchart from NASA Systems Engineering Handbook [3] emphasizing the change in a component's TRL resulting in a change of operating environment

Challenge 5 - Prioritization of technology development efforts

Description: Today's TRL assessments result in an evaluation of the maturity of dozens, or even hundreds of technologies. This takes time and effort. But the assessment is only information, not an action. How should the TRL information be interpreted? Should all low TRLs be addressed with equal effort? Should you start with the lowest TRL? Published TRL guidance lacks procedures for how to interpret and act on the TRL information.

Evidence: A popular current method for interpreting all of the component TRL information is called the "weakest link", where the lowest TRLs are "rolled up" the system hierarchy. For example, a subsystem's TRL is the lowest TRL of any component in the subsystem, and the system's overall TRL is the lowest TRL of any subsystem.

The weakest link method brings the lagging TRL component to the immediate attention of the manager, as explained by a director at Raytheon: "It quickly draws a red line around the low TRLs and suggests the program manager has to put some resources against them to fix it." But the weakest link is not perfect, since not all low TRL technologies should be the target of development focus, as explained by an engineering manager at BP: "You need new choke-and-kill outlet valves, which haven't been tested subsea, so they're at TRL 3. That puts the whole blowout preventer at TRL 3. Management gets a minor stroke. But you could just put the valves in a hyperbaric chamber and move those up to TRL 5 very quickly, without much effort." TRL users are not satisfied with the weakest link method.

Challenge 6 - System readiness

Description: Although the raw TRL information for each component is valuable to the developers as is, some have argued that there is a need for an overall system measure of technology maturity for the project. Ideally this system measure would indicate to the managers how the project as a whole is progressing from a technology development perspective. An effective means of displaying the portfolio of system readiness values is also lacking. This view would enable a readiness comparison to other projects in a portfolio.

Evidence: As described earlier, Sauser and his team recognized this need and developed the system readiness level [15]. Yet many practitioners shared the desire for an alternative system level measure, for example a NASA technologist explained that "SRL… everybody wants to use it. People would like to be able to characterize the maturity level of the system. Nobody has come up with anything that's useful yet." At John Deere, the systems engineering team previously explored using Sauser's SRL to measure the system readiness of one of their projects. The team ultimately decided it was not appropriate, explaining that “in the literature they were trying to calculate them by doing some matrix multiplication between the rating, which fell apart right away. […] We started working on the math and it didn’t work.”

Challenge 7 - Visualization

Description: Given that TRL assessment may involve hundreds of components, what is the most informative way in which to review and make decisions based on these data? Should the TRL data be aggregated or overlaid with other information? TRL assessments take a great deal of effort and attention, and it would be inefficient for the detail and quality of the TRL data to be lost in its visual representation.

Evidence: The most common means of sharing and reviewing TRL information is in a spreadsheet. As one engineer at Bombardier said, “we generate lists [of TRLs], and then pretty much use them listlessly.” NASA’s handbook shows a TRL spreadsheet with color coding to highlight low TRLs [3].

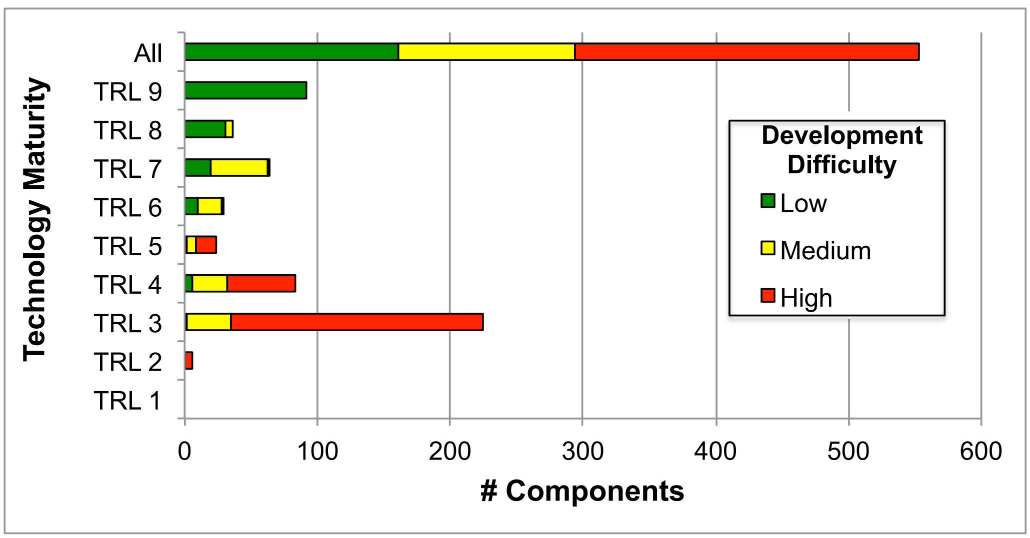

Figure 3: A summary of component TRL information for a complex project at BP, classified by both TRL and by development difficulty. (Note that the distribution of TRLs has been adjusted for confidentiality reasons.)

Figure 3 shows the type of bar chart that was provided to management of a BP project with over 500 component TRLs. Distinction between technologies as either low, medium or high development difficulty is included. This bar chart proved to be the most sophisticated visualization we discovered in our study.

B. Planning and Review

The challenges we have categorized as "planning and review" have to do with the decisions made during system development based on technology readiness progression. These challenges reflect the lack of integration of TRLs into the organization's existing business processes.

Challenge 8 – Aligning TRLs with system development gates

Description: In the introduction to this paper we described the practice of mapping the TRLs to the organization's standard system development process, with minimum acceptable TRLs at certain gates. Despite the adoption of this practice in industry, there has been little scholarly discussion of how to design such a mapping, how to determine the frequency of assessments, and how to determine appropriate minimum acceptable TRLs. Further, once this mapping is established, there is still a lack of understanding of the trade-offs and consequence of failure to achieve the goal TRL, which is a common scenario.

Evidence: A product design lead at Google described why establishing a mapping is a challenge for them, "Often times the product roadmap has a more regular cadence and a relatively short cycle. A lot of times the technology development that feeds into that, the cadence is not in synch." Thus the NASA or Department of Defense mapping is not appropriate for Google, but there is no real alternative.

At John Deere there is no formal standard mapping, despite the benefits being recognized: "[a mapping] helps people understand that they could not afford to be inventing new technology in a new product development program. [However, we were] on a new product development timeline and we were still doing technology invention."

Although formal mappings exist at some organizations, those that are new to TRLs encounter a challenge in establishing the appropriate mapping.

Challenge 9 - Waivers

Description: It is not uncommon when a decision gate is reached, that one or more technologies have not reached the minimum acceptable TRL. At that point, there are options: cancel the project, delay the project's advancement to the next stage until the lagging technologies are matured, or acknowledge the risk associated with the lagging technology, and move on to the next phase. Some organizations call this last option "dispensation", while others refer to issuing a "waiver". In either case, there is no commonly practiced method for giving such dispensation.

Evidence: The waiver process at BP requires a dispensation report be submitted to management, then be signed by 15 different people, including engineers, the project lead, independent verifiers and vice presidents, in order for the waiver to be issued.

According to the Department of Defense Technology Readiness Assessment Guidance [6], the decision to issue a waiver is made by the milestone decision authority, based on the project manager's risk-mitigation plans. Part of the risk mitigation may involve attempting to work with the customer to relax program requirements, or the inclusion of an alternative more mature technology. No formal waiver procedures were found at the other organizations interviewed, despite the common scenario of low TRL at a decision gate.

Challenge 10 - Back-up plans

Description: When developing risky technology, developers and managers often identify and pursue alternative technologies and options. If one of the technologies fails to mature by a certain decision gate, or the requirements evolve outside of the technology's performance range, there is a chance the alternative technologies will be appropriate and available (and necessary). There is no standard vocabulary with which to discuss back-up plans with respect to technology development and risk, and the concept is not reflected in technology readiness assessment guidelines.

Evidence: In our interview, a NASA technology manager explained that "[having] the fallback or alternative path or plan B is a 'best practice' but not a requirement at NASA. Many projects don't develop [such] exit ramps." We consistently learned from engineers and managers in our interviews that keeping back-up plans in place is a basic strategy familiar to and implemented on many projects, however there is no formalized link between back-up plans decisions (e.g. what is the cost and acceptable risk level?), technology readiness, and reviews.

Challenge 11- Effort to progress

Description: The TRL is a metric that reflects the technology's current maturity at a particular moment in the development lifecycle. Practitioners have come to realize that this technology "snapshot" does not provide any information about achieving future TRLs. In particular, an understanding of the effort, time, and resources required to progress to the next TRL and subsequent TRLs is lacking, and could be informative when planning and evaluating technology options.

Evidence: At Google the product design lead had practical questions about effort to progress: "What do we need to do to get it to the next stage of readiness? What do we really need to do to really have it secured in our back pocket, and put it on the shelves." From an interview with NASA we learned that effort to progress metrics are not consistently used and there is no standard process within the agency. In fact, this is an issue that NASA has independently identified as something they would like to investigate and develop further.

Challenge 12 - Confidence to progress

Description: Closely related to the previous challenge of effort to progress is the need for a measure of likelihood of forward progress in the TRL scale. Such confidence measures are generally based on gut instinct, and lack a methodical assessment technique.

Evidence: A risk expert at Alstom explained to us that the organization's "innovation" team uses TRL information to inform the probability of technical success of a key technology when making long-term technology selection decisions. The TRL is one criteria used to determine the probability of technical success, plotted on the vertical axis of a risk-reward chart, as shown in Figure 4. This chart is the only example we came across of likelihood to progress being used in the organizations studied, despite the promise of such a measure. There lacks a clear formulation and application for confidence to progress.

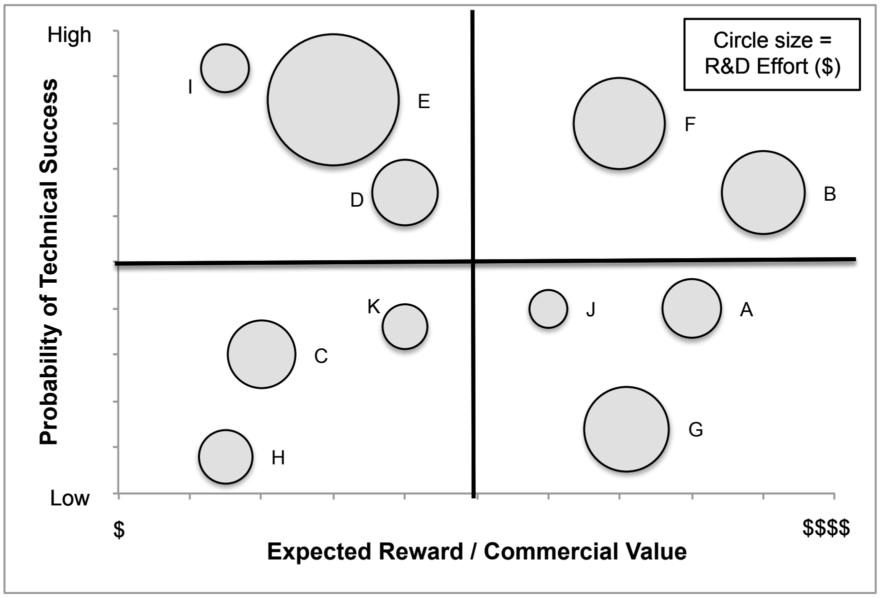

Figure 4: Representative example of a risk-reward chart as used at Alstom, demonstrating a use for the likelihood to progress measure. Each circle represents a prospective key technology.

Challenge 13 - Product roadmapping

Description: Product roadmaps are used as a planning tool to chart future versions of the product. This allows a shared understanding across the organization and can align decision making with regards to future product lines. Selecting which technologies to include in the product roadmap poses a complex challenge; the challenge is similar to that of alignment between the TRLs and the development process milestones, but it adds another dimension – not only do developers need to decide if technologies are mature enough for the product currently under development, but they need to think ahead to future product lines, and ensure that appropriate technologies that match the expected market and technical needs will be available for those products. When should an organization consider a technology ready for a future product in its roadmap?

Evidence: In the consumer electronics industry, it is common for a new product in a line to be launched within 12 months of its predecessor. New technologies need to be introduced to these products. But those technologies can take two or three years to develop. A product design lead at Google explained that "Right now it's pretty haphazard, where we're like 'ok that looks great, do you think it'll be ready by the time we do the next product?'", further stating that "we're looking for ways to validate when the technology is ready for dropping into the product roadmap to help us guide our development plans." The link between TRLs and roadmapping is not yet clear enough for actionable guidance.

C. Assessment Validity

The final category of challenges has to do with the quality of the TRL values achieved in a typical TRL assessment. Considering that critical development decisions are made on the basis of the assessed TRL, it is important that the values reached are valid and repeatable.

Challenge 14 - Subjectivity of the assessment

Description: The TRL is a 9-step scale with brief descriptions for each level. The levels are not perfectly distinct from one another, and for complex technologies it can be difficult for assessors to agree on precisely what level has been achieved. Assessors may have different expertise, experience and preconceptions. Since technologies can be highly complex, it is sometimes necessary for the technology lead (as the person best informed of the detail) to self-assess the technology readiness level. This creates an opportunity for bias due to optimism, competition for resources, avoidance of scrutiny, or merely awareness of their own sunk costs.

Evidence: An engineering manager from John Deere shared that when TRLs are used to inform technology selection decisions "inevitably the person who favors the technology will interpret the TRL higher than everybody else." An executive at BP summed up this challenge well, saying "if the three of us were working on a project together, do you think we'd assess all the equipment at the same TRL?"

Challenge 15 - Imprecision of the scale

Description: TRL definitions use the words "demonstration" and "validation" and "proven". These words can be interpreted to imply different deliverables. Given the complexity of today's technology, it is likely that a technology would be demonstrated in a series of tests. How should a developer deal with the common scenario where some tests are passed, and others are not? And is the test an operating test, or a stretch (worst-case scenario) test? What are the necessary test conditions and performance requirements? These are questions that are not addressed by current TRL guidelines.

Evidence: Alstom's various divisions follow at least three different versions of a customized TRL scale. This is perhaps due to the desire of those adopting the TRL scale for increased specificity of the language to suit the particular technology domain.

In the experience of a risk manager at Bombardier, the imprecision of the demonstration requirement inevitably leads to a discussion amongst the peer reviewers: "We go into a TRL review, and everyone comes in and has read the definition, but has not necessarily interpreted it in the same way. One might ask for a specific test to be done, while another says 'no no, we can just do a simulation."

Some developers, like a systems engineering manager at John Deere, understand that the generic TRL descriptions are not appropriate, explaining "If you're going to have an assessment and use this to make decisions, you're going to need criteria that are not only industry specific, but even product-line or product-type specific." Yet at this organization, there's no standard process for tailoring the TRL definitions, or formalizing the demonstration criteria.

IV. DISCUSSION

TRL-based technology assessment has been widely adopted across industry in the 40 years since its invention, proving its potential for sustainable positive impact. A modern investigation of TRL implementation challenges is important since the scope of systems and the market environments present today are significantly different than those faced by NASA in the 1970s. The 15 challenges identified in this work are driven by the increasing technical complexity of modern system design, growing incentives for TRL use, and practitioners' desire to expand the influence of TRLs to more decisions and organizational processes than were originally anticipated by NASA.

Although there is little previous work that systematically evaluates the state of TRL implementations, the limited evidence of shortcomings does agree with our findings. For instance, in a 2009 TRL retrospective, Mankins concluded with a short discussion of TRL challenges and directions. He identified the need for "practices and metrics that allow assessment of anticipated research and development uncertainty," the challenge we termed likelihood to progress. Further, he described the difficulty in achieving the right level of maturity across subsystems and components, a concept that we found to be decomposable and is therefore represented by multiple challenges in the system complexity category. Finally, Mankins emphasized the need for consistent assessment, which we consider to be congruent with our challenges related to the validity of assessment.

A. Improvement Opportunities

Each of the challenges described in this paper can be seen as an opportunity for improvement, and we believe that potential solutions may exist for all 15 challenges. Some solutions are straightforward, and may be solved by the simple sharing and implementation of best practices across industries; other solutions are not obvious, and present the opportunity for academic research. In this section we will outline our preliminary thoughts on potential directions forward.

1. System Complexity

Some of the solutions in the system complexity category can be addressed by taking advantage of systems engineering and architectural knowledge. A better understanding of the impact of integration immaturity effects (*challenge 1*) on the system architecture could be achieved through a cascade model, similar to that used in studies of change prediction in product redesign and customization [24]. In such a model, the cost and schedule risk of a low-TRL component could be considered to affect the cost and schedule risk of its neighboring components in the architecture because of shared interfaces. This solution could be informed by the work of Smaling and De Weck and their concept of invasiveness visualized in the delta design structure matrix [25]. Such a model could require advancements to the interface readiness concept (*challenge 2*). We believe that there remains an opportunity to explore alternatives to Sauser's integration readiness levels.

With regards to scope of the TRL assessment (*challenge 3*), we envision that a system architecture model will allow consideration of the trade-off of completeness (high cost) versus efficiency (risk of omitted issue) represented by the two approaches to TRL assessment. This could lead to insight regarding the conditions under which each method is appropriate.

An opportunity exists to develop a standard approach for TRL assessments when the system experiences an incremental change of component or environment (*challenge 4*). Given that best practices guidance is available at NASA for this concern, we believe this challenge may be straightforward to address. In terms of prioritization of development efforts (*challenge 5*), an alternative to the weakest link method of TRL roll up could be created by extending the TRL assessment to include some information regarding the risk of unsuccessful development of the technology. This information could be derived in a manner similar to the effort and likelihood to progress measures. Guidance for project managers on how to interpret and act on the results of a technology readiness assessment is needed.

Regarding the system readiness metric (*challenge 6*), the need for a simple measure to describe the system maturity should first be questioned. Is it reasonable that one simple number can appropriately and usefully describe a complex system? We are doubtful. This challenge may instead be best addressed with improved visualization (*challenge 7*) of all the component TRL information (and maybe IRL information), since a concise visualization of the full system technology readiness could provide the manager with an understanding of the system's overall maturity. Existing academic TRL visualization forms have centered around the product architecture, for example using the design structure matrix [17], and we believe this is a useful view. Improvements to this view could include additional information about development risk (for example, development difficulty as shown in the bar chart of Figure 3).

2. Planning and Review

Although identified as a challenge, industry leaders have successfully mapped their system development lifecycles to the TRLs (*challenge 8*). Learning from this best practice should provide insight related to the trade-offs between selecting immature technologies (with high technical potential at high development risk) or mature proven technologies (with limited technical potential at low development risk). A dynamic model would allow us to further our understanding of the technology-related choices made at decision gates. This model could encompass the concepts of waivers and back-up plans (*challenges 9 and 10*) and inform the choice between dispensations, delay, and cancel. A deeper understanding of what factors should be more highly considered by managers in their decision to grant or reject a waiver would be valuable. For example, should a waiver more readily be granted if a solid back-up plan has been identified? Although on its surface, the formalization of the link between back-up plans and TRLs appears clear (a low TRL technology should have a clear back-up plan), the issue has some interesting details that should be considered. Shishko et al. have explored the use of real options in technology decision making, which should inform future work on this challenge [26].

Such a model could benefit from reliable procedures and metrics for effort and likelihood to progress (*challenges 11 and 12*). A probabilistic model which considers confidence in the development success of the technology may help us better understand the answers to questions such as: When should technologies be developed in parallel as alternatives versus when should focus be placed on one major technology with a step-down technology identified? What should trigger the move to the step-down position? Present in the aerospace guidance, but not commonly practiced is the advancement degree of difficulty (referred to by NASA and Bilbro [3], [4]) or the research and development degree of difficulty (referred to by Mankins [12]). Neither of these measures is popular in industry and thus we believe there exists the opportunity to establish instead the confidence and effort to progress.

The TRL provides information on the current state of the technology maturity; combined with a measure of effort or likelihood, the TRL could be a powerful tool for selecting technologies to complete the product roadmap (*challenge 13*). It could be fruitful to use the TRL lens to study the work of Phaal et al. and Kostoff and Scaller, which provide thorough discussions of technology road-mapping practice [27], [28].

3. Validity of Assessment

Given that the concern for the quality of the TRL information is fundamental to all subsequent analysis, evaluations, and decisions, we believe the validity of assessment challenges present an important opportunity to study the repeatability and reproducibility of the TRL assessment. In addressing subjectivity concerns (*challenge 14*), there exists some preliminary work that explores the application of modern computing advances – such as computer document classification and big data – in assigning TRLs in a more automated way [29], [30]. The imprecision of the TRL scale (*challenge 15*) relates closely to the relationship between testing specifications and TRLs. A clearer understanding of this relationship may help to address some of the other challenges presented in this paper – a new environment or a new component will surely require different testing specifications, and a revisit of some TRLs. There is an opportunity to study the ad-hoc conversations and negotiations that occur while assigning TRLs at assessment meetings. This knowledge could inform the design of procedures or guidelines for tailoring the TRLs to match a specific project's testing specifications.

B. Observed Trends in TRL Process Maturity and Adoption

The variability in observed TRL challenges speaks to the learning curve associated with TRL implementation maturity. Even the most accomplished practitioners such as NASA are not fully satisfied with their implementations and are trying to improve their processes.

We plotted the number of challenges identified at each organization versus the length of time TRL processes have been used at each organization. We observed the resulting trend increase as organizations gained experience, and then decrease as practices gained maturity. NASA, for example, having used TRLs the longest, recognized being challenged by only a few concepts. Similarly, Google, new to TRL implementation, recognized a limited number of challenges. On the other hand, organizations like John Deere and Bombardier that have adopted TRL practices but are in the midst of making improvements identified the most challenges.

The trend aligns with expectations from learning theory, where we expect to see S-curve dynamics in similar applications such as gaining skill competence, process improvement, or quality management [31]–[33].

Our work has not focused on the adoption and acceptance of TRLs within organizations, however given the increasing number of organizations that are adopting TRLs, we believe such an investigation could reveal even further challenges. Additionally, a better understanding of the way in which TRL practices spread from industry to industry could help to anticipate future use-cases and challenges. Such studies have been done for similar processes in the past. In the context of software development maturity process innovations, Zeikowitz explores mechanisms for adoption of software innovation processes within NASA and across a segment in the aerospace industry [34]. Further relevant work is Repenning's study of innovation implementation (one could consider the TRL scale as such an innovation), which includes a dynamic model of the long-term adoption of innovations [35]. Anecdotally, we learned in our interviews that TRL exposure occurred in some of the organizations studied because of contractual obligation, regulatory requirement, cross-industry employee recruitment, or executive education programs.

V. CONCLUSIONS

With this paper, we provide a broad based discussion of the state-of-the-art in TRL practices and then identify 15 challenges associated with TRL implementations. We studied these TRL implementations through semi-structured interviews with employees from seven different organizations from diverse industries and by examining documentation collected from industry standards and organizational guidelines related to technology development and demonstration. Although some organizations are lagging behind best practices, even the most advanced TRL users face difficulty related to three categories of challenges: system complexity, planning and review, and assessment validity. With the increasing adoption of the TRL scale by system development firms further and further from the initial NASA context, it is important to evaluate the limitations and opportunities for improvement of this 40-year-old scale.

This paper provides practical insight to technology developers and managers in the form of increased awareness to potential pitfalls, a discussion of industry-spanning best practices, and suggestions for process improvement.

Progress towards addressing the opportunities presented in this work shows promise of a major positive impact on decision making, and resultant performance outcomes, regarding the injection, development, and integration of technologies in complex engineering projects.

REFERENCES

[1] S. Sadin, F. Povinelli, and R. Rosen, "The NASA Technology Push Towards Future Space Mission Systems," *Acta Astronaut.*, vol. 20, pp. 73–77, 1989.

[2] J. C. Mankins, "Technology Readiness Levels: A White Paper," 1995. Available: http://www.hq.nasa.gov/office/codeq/trl/trl.pdf.

[3] NASA, "NASA Systems Engineering Handbook," Washington, D.C., 2007.

[4] J. W. Bilbro, "Systematic Assessment of the Program / Project Impacts of Technological Advancement and Insertion Revision A," 2007.

[5] Det Norske Veritas, "Recommended Practice DNV-RP-A203 Qualification of New Technology," 2011.

[6] US Department of Defense, "Technology Readiness Assessment (TRA) Guidance," 2011. Available: http://www.acq.osd.mil/chieftechnologist/publications/docs/TRA2011.pdf.

[7] Homeland Security Institute, "Department of Homeland Security Science and Technology Readiness Level Calculator (ver 1.1) Final Report and User's Manual," 2009.

[8] United States Government Accountability Office, "Assessments of Selected Weapon Programs GAO-07-406SP," 2007.

[9] D. R. Katz, S. Sarkani, T. Mazzuchi, and E. H. Conrow, "The Relationship of Technology and Design Maturity to DoD Weapon System Cost Change and Schedule Change During Engineering and Manufacturing Development," *Syst. Eng.*, vol. 18, no. 1, pp. 1–15, 2015.

[10] America Makes, "About America Makes," 2015. [Online]. Available: https://americamakes.us.

[11] European Commission, "Horizon 2020: The EU Framework Programme for Research and Innovation," 2015. [Online]. Available: http://ec.europa.eu/programmes/horizon2020/en.

[12] J. C. Mankins, "Approaches to strategic research and technology (R&T) analysis and road mapping," *Acta Astronaut.*, vol. 51, no. 1–9, pp. 3–21, 2002.

[13] J. C. Mankins, “Technology readiness assessments: A retrospective,” *Acta Astronaut.*, vol. 65, no. 9–10, pp. 1216–1223, Nov. 2009.

[14] B. J. Sauser, R. Gove, E. Forbes, and J. E. Ramirez-Marquez, “Integration maturity metrics: Development of an integration readiness level,” *Inf. Knowl. Syst. Manag.*, vol. 9, pp. 17–46, 2010.

[15] B. J. Sauser, J. E. Ramirez-Marquez, D. Henry, and D. DiMarzio, “A system maturity index for the systems engineering life cycle,” *Int. J. Ind. Syst. Eng.*, vol. 3, no. 6, pp. 673–691, 2008.

[16] H. Jimenez and D. N. Mavris, “Characterization of Technology Integration Based on Technology Readiness Levels,” *J. Aircr.*, vol. 51, no. 1, pp. 291–302, Jan. 2014.

[17] S. D. Eppinger and T. R. Browning, *Design Structure Matrix Methods and Applications*. Cambridge, MA: MIT Press, 2012.

[18] T. K. Brady, “Utilization of Dependency Structure Matrix Analysis to Assess Complex Project Designs,” in *ASME 2002 Design Engineering Technical Conferences and Computer and Information in Engineering Conference*, 2002, pp. 1–10.

[19] S. Yasseri, “Subsea system readiness level assessment,” *Underw. Technol. Int. J. Soc. Underw.*, vol. 31, no. 2, pp. 77–92, Mar. 2013.

[20] N. Azizian, S. Sarkani, and T. Mazzuchi, “A comprehensive review and analysis of maturity assessment approaches for improved decision support to achieve efficient defense acquisition,” in *Proceedings of the World Congress on Engineering and Computer Science*, 2009, vol. II.

[21] N. Azizian, T. Mazzuchi, S. Sarkani, and D. Rico, “A framework for evaluating technology readiness, system quality, and program performance of US DoD acquisitions,” *Syst. Eng.*, vol. 14, no. 4, pp. 410–427, 2011.

[22] S. L. Cornford and L. Sarsfield, “Quantitative methods for maturing and infusing advanced spacecraft technology,” *2004 IEEE Aerosp. Conf. Proc.*, pp. 663–681, 2004.

[23] A. Edmondson and S. McManus, “Methodological fit in management field research,” *Acad. Manag. Rev.*, vol. 32, no. 4, pp. 1155–1179, 2007.

[24] P. J. Clarkson, C. Simons, and C. Eckert, “Predicting Change Propagation in Complex Design,” *J. Mech. Des.*, vol. 126, no. 5, p. 788, 2004.

[25] R. Smaling and O. De Weck, “Assessing Risks and Opportunities of Technology Infusion in System Design,” *Syst. Eng.*, vol. 10, no. 1, pp. 1–25, 2006.

[26] R. Shishko, D. H. Ebbeler, and G. Fox, “NASA technology assessment using real options valuation,” *Syst. Eng.*, vol. 7, no. 1, pp. 1–13, 2004.

[27] R. Phaal, C. J. P. Farrukh, and D. R. Probert, “Technology roadmapping—A planning framework for evolution and revolution,” *Technol. Forecast. Soc. Change*, vol. 71, no. 1–2, pp. 5–26, Jan. 2004.

[28] R. N. Kostoff and R. R. Schaller, “Science and Technology Roadmaps,” *IEEE Trans. Eng. Manag.*, vol. 48, no. 2, pp. 132–143, 2001.

[29] B. Britt, M. Berry, M. Browne, M. Merrell, and J. Kolpack, “Document classification techniques for automated technology readiness level analysis,” *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, no. 4, pp. 675–680, 2008.

[30] S. Cunningham, “Big data and technology readiness levels,” *IEEE Eng. Manag. Rev.*, vol. 42, no. 1, pp. 8–9, Mar. 2014.

[31] L. Adams, “Learning a New Skill is Easier Said Than Done,” 2011. [Online]. Available: http://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/.

[32] W. I. Zangwill and P. B. Kantor, “Toward a Theory of Continuous Improvement and the Learning Curve,” *Manage. Sci.*, vol. 44, no. 7, pp. 910–920, Jul. 1998.

[33] C. D. Ittner, V. Nagar, and M. V. Rajan, “An Empirical Examination of Dynamic Quality-Based Learning Models,” *Manage. Sci.*, vol. 47, no. 4, pp. 563–578, Apr. 2001.

[34] M. V Zelkowitz, “Software Engineering Technology Infusion Within NASA,” *IEEE Trans. Eng. Manag.*, vol. 43, no. 3, pp. 250–261, 1996.

[35] N. P. Repenning, “A Simulation-Based Approach to Understanding the Dynamics of Innovation Implementation,” *Organ. Sci.*, vol. 13, no. 2, pp. 109–127, Apr. 2002.